

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:55:01 ; Search time 1856.32 Seconds

(without alignments)
14423.488 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920

Sequence: 1 tgcacacatgcacaaacaa.....gtagtattataaaaaaaa 920

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vt:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vt:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Desc:ption
1	920	100.0	920	6 A16257	16257 Synthetic R
2	920	100.0	920	6 A16258	A16258 Synthetic R
3	920	100.0	920	6 AR080406	AR080406 Sequence
4	920	100.0	920	6 AR092530	AR092530 Sequence
5	920	100.0	920	6 AR122885	AR122885 Sequence
6	920	100.0	920	6 AR123540	AR123540 Sequence
7	920	100.0	920	6 AR148357	AR148357 Sequence
8	911.8	99.1	923	14 HRSRNG	X17085 Human respl
9	909.2	98.8	922	14 HRSVGL16	Z33429 Human respl
10	901.6	98.0	917	14 RSHGLYG	M17212 Human respl
11	894	97.2	894	6 AR080424	AR080424 Sequence
12	894	97.2	894	6 AR092548	AR092548 Sequence
13	853.4	92.8	8510	14 RSHICE	M17486 Human respl
14	853.4	92.8	15222	6 AR093219	AR093219 Sequence
15	853.4	92.8	15222	14 HRU50362	U50362 Human respl
16	853.4	92.8	15222	14 HRU50363	U50363 Human respl
17	853.4	92.8	15222	14 HRU63644	U63644 Human respl
18	853.4	92.8	15222	14 RSHSEO	M74568 Human respl
19	853.4	92.8	15223	6 AR089137	AR089137 Sequence
20	853.4	92.8	15223	14 AF035006	AF035006 Human res
21	848.2	92.2	914	6 PARSENG	X03149 Respiratory
22	848.2	92.2	918	6 AX047075	AX047075 Sequence
23	837.8	91.1	897	6 AX339011	AX339011 Sequence
24	837.2	91.0	15190	14 HRU39662	U39662 Human respl
25	835.6	90.8	15191	14 RSU39661	AF065406 Human res
26	820.4	89.2	918	14 AF065406	AF065406 Human res
27	814	88.5	918	14 AF065405	Z33410 Human respl
28	811.6	88.2	922	14 HRSVGL1	Z33412 Human respl
29	811.6	88.2	922	14 HRSVGL1	Z33412 Human respl
30	810	88.0	922	14 HRSVGL1	Z33412 Human respl
31	810	88.0	922	14 HRSVGL1	Z33412 Human respl
32	810	88.0	922	14 HRSVGL1	Z33412 Human respl
33	810	88.0	922	14 HRSVGL1	Z33412 Human respl
34	808.4	87.9	922	14 HRSVGL1	Z33412 Human respl
35	808.4	87.9	922	14 HRSVGL1	Z33412 Human respl
36	807.2	87.7	922	14 HRSVGL1	Z33412 Human respl
37	806.8	87.7	922	14 HRSVGL1	Z33412 Human respl
38	806.8	87.7	922	14 HRSVGL1	Z33412 Human respl
39	805.2	87.5	922	14 AF065256	AF065256 Human res
40	805.2	87.5	922	14 HRSVGL1	Z33423 Human respl
41	805.2	87.5	922	14 HRSVGL1	Z33423 Human respl
42	805.2	87.5	922	14 HRSVGL1	Z33423 Human respl
43	803.6	87.3	922	14 AF065255	AF065255 Human res
44	803.6	87.3	922	14 AF065258	AF065258 Human res
45	803.6	87.3	922	14 HRSVGL1	Z33430 Human respl

ALIGNMENTS

RESULT 1
LOCUS A16257 920 bp DNA
DEFINITION Synthetic RSV G gene (seq ID No: 7).
ACCESSION A16257
VERSION A16257.1 GI:640933
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequences.
FEATURES
source
1..920
/organism="synthetic construct"
/db_xref="taxon:32630"
8..904
gene

Pred. No. is the number of results predicted by chance to have a

QY	601	CACCCGAAAGCTCAACAAAAAACCACTTCAAGCAACCAAAAAAGATCTTAAACTCA	660
Db	320	CACCAACCAAGCTCAACAAAAAACCACTTCAAGCAACCAAAAAAGATCTTAAACTCA	261
QY	661	AACCACTAAACCAAGAGATACCCACCACCAAGGCCACAGAGAGCCAAACCATCAAC	720
Db	260	AACCACTAAACCAAGAGAGATACCCACCACCAAGGCCACAGAGAGCCAAACCATCAAC	201
QY	721	CACCAAAACCAACATCAACATACACCTCTCAACCAACAACACACAGGAAATCCAAACT	780
Db	200	CACCAAAACCAACATCAACATACACCTCTCAACCAACAACACACAGGAAATCCAAACT	141
QY	781	CACAAGTCAAAATGGAACCTTCACACTCAACCTCTCTCCGAAGCAATCTTAAGCCCTTCA	840
Db	140	CACAAGTCAAAATGGAACCTTCACACTCAACCTCTCTCCGAAGCAATCTTAAGCCCTTCA	81
QY	841	AGTCTCCGAACATCCCGAGCACCCATCAACACCCATCTCCACCCACACACAACAGGCA	900
Db	80	AGTCTCCGAACATCTCCGAGCACCCATCAACACCCATCTCCACCCACACACAACAGGCA	21
QY	901	GTAGTATATAAAAAAAA 920	
Db	20	GTAGTATATAAAAAAAA 1	

RESULT 3					
AR080406					
LOCUS	AR080406	920 bp	DNA	linear	PAT 31-AUG-2000
DEFINITION	Sequence 7 from patent US 5968776.				
ACCESSION	AR080406				
VERSION	AR080406.1	GI:10007141			
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				

REFERENCE	AUTHORS	TITLE
1 (pages 1 to 920)	Klein, M. H., Du, R.-P. and Ewaszshyn, M. E.	Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus

JOURNAL Patent: US 5968776-A 7 19-OCT-1999;
FEATURES Location/Qualifiers

```

source      1. 920
            /organism="unknown"
BASE COUNT 380 a 290 c 95 g 155 t
BRIGIN

```

Query Match	100.0%	Score 920;	DB 6;	Length 920;
Best Local Similarity	100.0%	Pred. No. 3.6e-190;		
Matches 920; Conservative	0;	Mismatches	0;	Gaps 0;

OY	1	TGCAAAATGTGCCAAAAACAAGACCAAGCGACGGCTAAGCACTAGAAAGACCTGGGA	60
Db	1	TGCAACATGTCCAAAAACAAGACCAAGCGACGGCTAAGCACTAGAAAGACCTGGGA	60
OY	61	CACCTCCAATCATTTATTATTCATATATCATGCGGCTTATATAGTTAAATCTTAAATCTGT	120
Db	61	CACCTCCAATCATTTATTATTCATATATCATGCGGCTTATATAGTTAAATCTTAAATCTGT	120
OY	121	AGCAACAATACATTTATTCATCTTGGCAATGATATATCTCAACTTCATTATATTTACAGC	180
Db	121	AGCAACAATACATTTATTCATCTTGGCAATGATATATCTCAACTTCATTATATTTACAGC	180
OY	181	CATCATTTTCATAGCCCTGGCAACCAAGTCAACATACCACTGCAATCATACAAAGA	240
Db	181	CATCATTTTCATAGCCCTGGCAACCAAGTCAACATACCACTGCAATCATACAAAGA	240
OY	241	TGCACAAGCAGATCAAGAACAACGCCAACAATACCTCACTAGAGATCCTCAGCTTGG	300
Db	241	TGCACAAGCAGATCAAGAACAACGCCAACAATACCTCACTAGAGATCCTCAGCTTGG	300
OY	301	AATTCAGCTTCCAAATCTGTGAATTTACATCAACAACCAACCACTACTAGCTTCAC	360

Db	301	AATCGCTTTCGCAATCTGTCTGAAATTACATGCACAAACCCACCACCTACTAGTCTTAAC	360
QY	361	AACACGAGAGTCAAGTCAAAACCTGCACCCACCAACAGTCAAGACTCAAAAACACAACAC	420
Db	361	AACACGAGAGTCAAGTCAAAACCTGCACCCACCAACAGTCAAGACTCAAAAACACAACAC	420
QY	421	AACCCAAACACACCCACGCAAGCCCATACAAAACAAGCCCAAAACCAACACAC	480
Db	421	AACCCAAACACACCCACGCAAGCCCATACAAAACAAGCCCAAAACCAACCAACAA	480
QY	481	ACCCAAATATGATTTTCGACTTCGAAAGTGTTAACTTTGACCCCTGACATATGCAGCA	540
Db	481	ACCCAAATATATATTTTCGACTTCGAAAGTGTTAACTTTGACCCCTGACATATGCAGCA	540
QY	541	CAATCCAACTGCTGGGCTATCTGCAAAAGAAATATCCAAACCAAAACACGAAAGAAAC	600
Db	541	CAATCCAACTGCTGGGCTATCTGCAAAAGAAATATCCAAACCAAAACACGAAAGAAAC	600
QY	601	CACGACCAAGCTTCAAAAAAACCACTTCAAGACAAACCAAAAAAGATCTTAAACCTCA	660
Db	601	CACGACCAAGCTTCAAAAAAACCACTTCAAGACAAACCAAAAAAGATCTTAAACCTCA	660
QY	661	AACCACTAAACCAAGGAGTACCCACACCAAGGCCACAGAAAGGCCAATCATACAC	720
Db	661	AACCACTAAACCAAGGAGTACCCACACCAAGGCCACAGAAAGGCCAATCATACAC	720
QY	721	CACCAAAAACAACATCCAACTACACTGCTCACCACCAACACACACAGAAATCCAAAAC	780
Db	721	CACCAAAAACAACATCCAACTACACTGCTCACCACCAACACACACAGAAATCCAAAAC	780
QY	781	CACAAGTCAATGGAACCTTTCACATCTCAACTCTCCGAAGGCATCTTAAGCCCTTCTCA	840
Db	781	CACAAGTCAATGGAACCTTTCACATCTCAACTCTCCGAAGGCATCTTAAGCCCTTCTCA	840
QY	841	AGTCTCCACAACATCCGACACCCATACAAACCTCATCTCCACCCAAACACAAACGCA	900
Db	841	AGTCTCCACAACATCCGACACCCATACAAACCTCATCTCCACCCAAACACAAACGCA	900
QY	901	GTAGTTATTAATAAAAAAAA 920	
Db	901	GTAGTTATTAATAAAAAAAA 920	

RESULT	4
AR092530	
LOCUS	AR092530
DEFINITION	Sequence from patent US 5998169.
ACCESSION	AR092530.1
VERSION	GI:10019284
	920 bp
	DNA
	linear
	PAT 08-SEP-2000

REFERENCE	AUTHORS	TITLE
Unclassified. 1 (bases 1 to 920)	Klein, M. H., Du, R.-P. and Ewasyschyn, M. E.	Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial

VIRUS
JOURNAL Patent: US 5998169-A 7 07-DEC-1999;
FEATURES location/Qualifiers

SOURCE	1: 920	/organism="unknown"		
BASE COUNT	380 a	290 c	95 g	155 t
ORIGIN				

Query Match	100.0%	Score 920;	DB 6;	Length 920;
Best Local Similarity	100.0%	Pred. No. 3.6e190;		
Matches 920; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 TGCACACATGTCMAAAACAAGGACCAACGACACCGCTAAGACACTGAGAAAGACCTGGGA 60

Db 1 TGCACACATGTCMAAAACAAGGACCAACGACACCGCTAAGACACTGAGAAAGACCTGGGA 60

QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120
|||||
Db 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120
QY 121 AGCAAAATACATTTATTCATTTCTGGCAATGATATTCACATTCCTTATATAC 180
|||||
Db 121 AGCAAAATACATTTATTCATTTCTGGCAATGATATTCACATTCCTTATATAC 180
QY 181 CATCATTTATTCATAGCTTCGGGCAAAACGCAAAAGTACACTTAACATCTATACAGA 240
|||||
Db 181 CATCATTTATTCATAGCTTCGGGCAAAACGCAAAAGTACACTTAACATCTATACAGA 240
QY 241 TGCACACAGCCAGATCAAGAACACACCCACATACCTCAGTACGAGTCTGAGCTTG 300
|||||
Db 241 TGCACACAGCCAGATCAAGAACACACCCACATACCTCAGTACGAGTCTGAGCTTG 300
QY 301 AATAGCTTCTCCATCTGCTGTAATTTACATCAACACGACCACTACTAGCTTCAAC 360
|||||
Db 301 AATAGCTTCTCCATCTGCTGTAATTTACATCAACACGACCACTACTAGCTTCAAC 360
QY 361 AACACAGAGATCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAAAACACACAC 420
|||||
Db 361 AACACAGAGATCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAAAACACACAC 420
QY 421 AACCCAAACACACACCCGCAAGCCCACTACAAAACAGCCCAAAACACACCAACAA 480
|||||
Db 421 AACCCAAACACACACCCGCAAGCCCACTACAAAACAGCCCAAAACACACCAACAA 480
QY 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGACCTGCGAGATATGACGCA 540
|||||
Db 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGACCTGCGAGATATGACGCA 540
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAAACCCAGAAAAAAC 600
|||||
Db 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAAACCCAGAAAAAAC 600
QY 601 CACACCAAGGCTCAAAAAAACCACTTCAAGACAGCAACAAAAAATCTCAAACTCA 660
|||||
Db 601 CACACCAAGGCTCAAAAAAACCACTTCAAGACAGCAACAAAAAATCTCAAACTCA 660
QY 661 AACCACTAAACCAAGAGATACCCACACCAACCCACAGAGAGGCAACCATCAAC 720
|||||
Db 661 AACCACTAAACCAAGAGATACCCACACCAACCCACAGAGAGGCAACCATCAAC 720
QY 721 CACCAAAACCAACATCACTACACTGCTGACCAACACCAAGGAAATCCAAACT 780
|||||
Db 721 CACCAAAACCAACATCACTACACTGCTGACCAACACCAAGGAAATCCAAACT 780
QY 781 CACAAGTCAAAATGGAACCTTCCACACCTCTCCGAGAGGCAATTAAGCCCTTCTCA 840
|||||
Db 781 CACAAGTCAAAATGGAACCTTCCACACCTCTCCGAGAGGCAATTAAGCCCTTCTCA 840
QY 841 AGTCTCACAACATCGAGGACCCATCAACCTCTCTCAACCCCAACCAACAGCCCA 900
|||||
Db 841 AGTCTCACAACATCGAGGACCCATCAACCTCTCTCAACCCCAACCAACAGCCCA 900
QY 901 GTAGTTATTTAAAAAAA 920
|||||
Db 901 GTAGTTATTTAAAAAAA 920

RESULT 5
ARI22885 920 bp DNA linear PAT 16-MAY-2001
LOCUS
DEFINITION Sequence 7 from patent US 6168786.
ACCESSION ARI22885
VERSION ARI22885.1 GI:14107851
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein, M.H., Du, R.-P. and Ewasysbyn, M.E.

TITLE Chimeric immunogens
JOURNAL Patent: US 6168786-A 7 02-JAN-2001;
FEATURES location/Qualifiers
source 1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3,6e-190;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGCAAACATGTCACAAAACAGAGCAACGACCGCTAAGACACTAGAAAAAGCCTGGGA 60
|||||
Db 1 TGCAAACATGTCACAAAACAGAGCAACGACCGCTAAGACACTAGAAAAAGCCTGGGA 60
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120
|||||
Db 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTGT 120
QY 121 AGCAAAATACATTTATTCATTTCTGGCAATGATATTCACATTCCTTATATAC 180
|||||
Db 121 AGCAAAATACATTTATTCATTTCTGGCAATGATATTCACATTCCTTATATAC 180
QY 181 CATCATTTATTCATAGCTTCGGGCAAAACGCAAAAGTACACTTAACATCTATACAGA 240
|||||
Db 181 CATCATTTATTCATAGCTTCGGGCAAAACGCAAAAGTACACTTAACATCTATACAGA 240
QY 241 TGCACACAGCCAGATCAAGAACACACCCACATACCTCAGTACGAGTCTGAGCTTG 300
|||||
Db 241 TGCACACAGCCAGATCAAGAACACACCCACATACCTCAGTACGAGTCTGAGCTTG 300
QY 301 AATAGCTTCTCCATCTGCTGTAATTTACATCAACACGACCACTACTAGCTTCAAC 360
|||||
Db 301 AATAGCTTCTCCATCTGCTGTAATTTACATCAACACGACCACTACTAGCTTCAAC 360
QY 361 AACACAGAGATCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAAAACACACAC 420
|||||
Db 361 AACACAGAGATCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAAAACACACAC 420
QY 421 AACCCAAACACACACCCGCAAGCCCACTACAAAACAGCCCAAAACACACCAACAA 480
|||||
Db 421 AACCCAAACACACACCCGCAAGCCCACTACAAAACAGCCCAAAACACACCAACAA 480
QY 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGACCTGCGAGATATGACGCA 540
|||||
Db 481 ACCCAATATGATTTTCACTTCGAAGTGTAACTTTGACCTGCGAGATATGACGCA 540
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAAACCCAGAAAAAAC 600
|||||
Db 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAACAAAAAACCCAGAAAAAAC 600
QY 601 CACACCAAGGCTCAAAAAAACCACTTCAAGACAGCAACAAAAAATCTCAAACTCA 660
|||||
Db 601 CACACCAAGGCTCAAAAAAACCACTTCAAGACAGCAACAAAAAATCTCAAACTCA 660
QY 661 AACCACTAAACCAAGAGATACCCACACCAACCCACAGAGAGGCAACCATCAAC 720
|||||
Db 661 AACCACTAAACCAAGAGATACCCACACCAACCCACAGAGAGGCAACCATCAAC 720
QY 721 CACCAAAACCAACATCACTACACTGCTGACCAACACCAAGGAAATCCAAACT 780
|||||
Db 721 CACCAAAACCAACATCACTACACTGCTGACCAACACCAAGGAAATCCAAACT 780
QY 781 CACAAGTCAAAATGGAACCTTCCACACCTCTCCGAGAGGCAATTAAGCCCTTCTCA 840
|||||
Db 781 CACAAGTCAAAATGGAACCTTCCACACCTCTCCGAGAGGCAATTAAGCCCTTCTCA 840
QY 841 AGTCTCACAACATCGAGGACCCATCAACCTCTCTCAACCCCAACCAACAGCCCA 900
|||||
Db 841 AGTCTCACAACATCGAGGACCCATCAACCTCTCTCAACCCCAACCAACAGCCCA 900
QY 901 GTAGTTATTTAAAAAAA 920
|||||


```
Db 901 GTAGTTATTAAAAAAA 920
|||||
RESULT 6
LOCUS AR123540 920 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 7 from patent US 6171783.
ACCESSION AR123540
VERSION AR123540.1 GI:14108901
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE Infection detection method using chimeric protein
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;
FEATURES
source location/Qualifiers
1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.6e-190;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAAACATGTCCAAAAACAGACCAACGCCGCTAAGACTAGAGAAAAAGCCTGGGA 60
|||||
Db 1 TGCAAACATGTCCAAAAACAGACCAACGCCGCTAAGACTAGAGAAAAAGCCTGGGA 60
QY 61 CACTCTCATCTATTATTATTCATATCGGGCTTATATAGTTAAATCTTAATCTGT 120
|||||
Db 61 CACTCTCATCTATTATTATTCATATCGGGCTTATATAGTTAAATCTTAATCTGT 120
QY 121 AGCACAATCACATTATTCATTGTGGCAATGATATCTCAACTTCACTTATTAATTCAGC 180
|||||
Db 121 AGCACAATCACATTATTCATTGTGGCAATGATATCTCAACTTCACTTATTAATTCAGC 180
QY 121 CATTCAATTCATAGGCTGGGCAACCAAGCTCACTAATCAATGCAATCATCAAGA 240
|||||
Db 181 CATTCAATTCATAGGCTGGGCAACCAAGCTCACTAATCAATGCAATCATCAAGA 240
QY 181 CATCATATTTCATAGGCTGGGCAACCAAGCTCACTAATCAATGCAATCATCAAGA 240
|||||
Db 181 CATCATATTTCATAGGCTGGGCAACCAAGCTCACTAATCAATGCAATCATCAAGA 240
QY 241 TGCACAAGCCGATCAAGAACACACACCCCAATCACTCACTCGATCTGAGTCTGAGCTTGG 300
|||||
Db 241 TGCACAAGCCGATCAAGAACACACACCCCAATCACTCACTCGATCTGAGTCTGAGCTTGG 300
QY 301 AATCAGCTTCTGCAATCTGTCTGAATTTACATCACAAAACCAACCATCTAGCTTCAAC 360
|||||
Db 301 AATCAGCTTCTGCAATCTGTCTGAATTTACATCACAAAACCAACCATCTAGCTTCAAC 360
QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATTAATAAACACAGCAAC 420
|||||
Db 361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATTAATAAACACAGCAAC 420
QY 421 AACCCAAACAAACCCAGACGCCACTACAAAACCAAGCCCAAAACCAACCAACCAAA 480
|||||
Db 421 AACCCAAACAAACCCAGACGCCACTACAAAACCAAGCCCAAAACCAACCAACCAAA 480
QY 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGACCTGACGACATATGAGCAA 540
|||||
Db 481 ACCCAATATGATTTTCACTTGAAGTGTAACTTTGACCTGACGACATATGAGCAA 540
QY 541 CAATCAACCTGCTGGGTATCTGCAAAAAGATACAAAACCAAAACAGGAAAGAAAAC 600
|||||
Db 541 CAATCAACCTGCTGGGTATCTGCAAAAAGATACAAAACCAAAACAGGAAAGAAAAC 600
QY 601 CACCAACAAGCCTACAAAAAACCAAGCTTCAAGACACCAAAAAAGATCTCCAACCTCA 660
|||||
Db 601 CACCAACAAGCCTACAAAAAACCAAGCTTCAAGACACCAAAAAAGATCTCCAACCTCA 660
QY 661 AACCACTAAACCAAGAAATACCAACCAAGCCCAAGAGCCACAGAGAGCAACCATCAAC 720
|||||
```

```
Db 661 AACCACTAAACCAAGAAATACCAACCAAGCCCAAGAGCCACAGAAAGGCAATCAACAC 720
|||||
QY 721 CACCAAAACAACATCAACACTACACTGCTCACCACACACACAGAAATCCAAACT 780
|||||
Db 721 CACCAAAACAACATCAACACTACACTGCTCACCACACACAGAAATCCAAACT 780
QY 781 CACAAGTCAAAATGGAACCTTCACATCACTCCTCGAAGGCAATCTAGCCCTCTCA 840
|||||
Db 781 CACAAGTCAAAATGGAACCTTCACATCACTCCTCGAAGGCAATCTAGCCCTCTCA 840
QY 841 AGTCTCCACAACATCCGAGACCCATCACAAACCTCATCTCCACCAACCAACAGCCCA 900
|||||
Db 841 AGTCTCCACAACATCCGAGACCCATCACAAACCTCATCTCCACCAACCAACAGCCCA 900
QY 901 GTAGTTATTAAAAAAA 920
|||||
Db 901 GTAGTTATTAAAAAAA 920
|||||
RESULT 7
LOCUS AR148357 920 bp DNA linear PAT 08-AUG-2001
DEFINITION Sequence 7 from patent US 6225091.
ACCESSION AR148357
VERSION AR148357.1 GI:15112447
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers
protection against parainfluenza virus and respiratory syncytial
virus
JOURNAL Patent: US 6225091-A 7 01-MAY-2001;
FEATURES
source location/Qualifiers
1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 920; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.6e-190;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCAAACATGTCCAAAAACAGACCAACGCCGCTAAGACTAGAGAAAAAGCCTGGGA 60
|||||
Db 1 TGCAAACATGTCCAAAAACAGACCAACGCCGCTAAGACTAGAGAAAAAGCCTGGGA 60
QY 61 CACTCTCATCTATTATTATTCATATCGGGCTTATATAGTTAAATCTTAATCTGT 120
|||||
Db 61 CACTCTCATCTATTATTATTCATATCGGGCTTATATAGTTAAATCTTAATCTGT 120
QY 121 AGCAAAATCAATATTCATCTGGAATGATTAATCTCACTTATAATTACAGC 180
|||||
Db 121 AGCAAAATCAATATTCATCTGGAATGATTAATCTCACTTATAATTACAGC 180
QY 181 CATCATATTTCATAGCCTGGGCAAAACCAAGTCACTAATCAAGTCAATCTTACAAGA 240
|||||
Db 181 CATCATATTTCATAGCCTGGGCAAAACCAAGTCACTAATCAAGTCAATCTTACAAGA 240
QY 241 TGCACAAGCCGATCAAGAACACACACCCCAATCACTCACTCGATCTGAGTCTGAGCTTGG 300
|||||
Db 241 TGCACAAGCCGATCAAGAACACACACCCCAATCACTCACTCGATCTGAGTCTGAGCTTGG 300
QY 301 AATCAGCTTCTGCAATCTGTCTGAATTTACATCAAAAACCAACCATCTAGCTTCAAC 360
|||||
Db 301 AATCAGCTTCTGCAATCTGTCTGAATTTACATCAAAAACCAACCATCTAGCTTCAAC 360
QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATTAATAAACACAGCAAC 420
|||||
Db 361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATTAATAAACACAGCAAC 420
|||||
```


[illegible]

QY	241	TTGAACCAAGCCAGATCAAGAACACACACACCCCAACATACCTCACTCGAGGATCCGACCTGG	300
Db	249	TTGCAACAAGCCAGATCAAGAACACACACCCCAACATACCTCACTCGAGGATCCGACCTGG	308
QY	301	AATGAGCTTCTCCATCTGTCTGTAATTAACATACACMAACCAACACCATTAAGTTCAAC	360
Db	309	AATGAGCTTCTCCATCTGTCTGTAATTAACATACACMAACCAACACCATTAAGTTCAAC	368
QY	361	AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGACTAAAAACATACCAAC	420
Db	369	AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGACTAAAAACATACCAAC	428
QY	421	AACCCAAACACACACCAGCAAGCCCACTACAAAACAGCCCAAAACCAACCAACAA	480
Db	429	AACCCAAACACACACCAGCAAGCCCACTACAAAACAGCCCAAAACCAACCAACAA	488
QY	481	ACCCAAATATGATTTTCACCTTGAAGTGTATTAACCTTTACCTGACACATATGACGAA	540
Db	489	ACCCAAATATGATTTTCACCTTGAAGTGTATTAACCTTTACCTGACACATATGACGAA	548
QY	541	CAATCCAACTGTCTGTGGCTATCTGCAAAAAGATACCAACAAAAAACCAAGAAATAAC	600
Db	549	CAATCCAACTGTCTGTGGCTATCTGCAAAAAGATACCAACAAAAAACCAAGAAATAAC	608
QY	601	CACCAACAGCTTACAAAAAAACCAACCTTGAAGCAACCAAAAAAGATTCACAACTCA	660
Db	609	CACCAACAGCTTACAAAAAAACCAACCTTGAAGCAACCAAAAAAGATTCACAACTCA	668
QY	661	AACCACTAAACCAAGAAAGTACCCACCAACCAAGCCCAAGAAAGCCACCATACACAC	720
Db	669	AACCACTAAACCAAGAAAGTACCCACCAACCAAGCCCAAGAAAGCCACCATACACAC	728
QY	721	CACCAAAACCAACATACACACTACACTGTCTACCAACACACACAGGAAATCCAAACT	780
Db	729	CACCAAAACCAACATACACACTACACTGTCTACCAACACACACAGGAAATCCAAACT	788
QY	781	CACAGTCAAAATGCAAACTTCCACTCAACTCTCCCAAGCAAGCAATTAATCTTCACA	840
Db	789	CACAGTCAAAATGCAAACTTCCACTCAACTCTCCCAAGCAAGCAATTAATCTTCACA	848
QY	841	AGTCTCCACACATCCGAGCACCCATCCACAAACCTCATCTCCACCCCAACACACGCGCA	900
Db	849	AGTCTCCACACATCCGAGCACCCATCCACAAACCTCATCTCCACCCCAACACACGCGCA	908
QY	901	GTAGTTATTTAAAA 914	
Db	909	GTAGTTATTTAAAA 922	
RESULT 10			
* RSHLXG			
LOCUS	Human respiratory syncytial virus (subgroup A) attachment protein		
DEFINITION	(g) mRNA, complete cds.		
ACCESSION	M17212.1 GI:333940		
VERSION	M17212		
KEYWORDS	attachment glycoprotein; surface glycoprotein.		
SOURCE	Human respiratory syncytial virus (subgroup A, strain Long), cDNA to viral RNA, clones pLF63, pLB69, and pA3.		
ORGANISM	Human respiratory syncytial virus		
REFERENCE	Viruses; ssRNA negative-strand virus; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.		
AUTHORS	1 (bases 1 to 917)		
TITLE	Johnson, P.R., Spriggs, M.K., Olmsted, R.A. and Collins, P.L.		
JOURNAL	The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins		
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 84 (16), 5625-5629 (1987)		
PUBMED	87289657		
COMMENT	The exact 5' end of Long G mRNA was not determined.		
FEATURES	Location/Qualifiers		

source 1. .917
/organism="Human respiratory syncytial virus"
/db_xref="taxon:11250"
16. .912
CDS
/note="attachement glycoprotein (G)"
/codon_start=1
/protein_id="AA47411.1"
/db_xref="GI:333941"
/translation="MSKNDQPTAKLEKTDWLNHLFISGLYKLNKLSINQITLS
ILAMISLILITALEIFASANKVLTLLTIDATSOIKNTPTLTODPOLGISFS
NLSITISOTTTILASTPGVSNLOPTVTKTKTODOPSKPTKORONKRPKNP
NDHFPEVNFVPSICSNPNPCMAICKRIPNKPGRKTTTKPKFTTKDKHNP
TTKREKPTTKPBEPTINTKINILITLNTGTPNKLTSOMETFSHSTSEGNLSP
SQVSTSEHPSPPSPNTTRO"
BASE COUNT 375 a 290 c 97 g 154 t 1 others
ORIGIN
Query Match 98.0%; Score 901.6; DB 14; Length 917;
Best Local Similarity 99.6%; Pred. No. 3.6e-186;
Matches 904; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 2 GCAACATGTCACAAAGAGACCAACGACGCTAGACACTAGAAAAGACCTGGAC 61
DB 10 GCAACATGTCACAAAGAGACCAACGACGCTAGACACTAGAAAAGACCTGGAC 69
QY 62 ACTCTCAATCATTTATTTATTCATTCATCGGGCTTATATAGTTAAATCTTAACTGTA 121
DB 70 ACTCTCAATCATTTATTTATTCATTCATCGGGCTTATATAGTTAAATCTTAACTGTA 129
QY 122 GCACAAATCACAATTCATTCATTCGCAATGATATTCACATTCATTAATTAATACGCC 181
DB 130 GCACAAATCACAATTCATTCATTCGCAATGATATTCACATTCATTAATTAATACGCC 189
QY 182 ATCATATTCATAGCTTCGGCAACACCAAGTCACATTAACACTGCAATCATACAGAT 241
DB 190 ATCATATTCATAGCTTCGGCAACACCAAGTCACATTAACACTGCAATCATACAGAT 249
QY 242 GCACAAAGCCGATCAAGAACACCAACCCCAATACCTGACTGAGATCTGACGTTGGA 301
DB 250 GCACAAAGCCGATCAAGAACACCAACCCCAATACCTGACTGAGATCTGACGTTGGA 309
QY 302 ATCAGCTTCTCAATCTGTCTGAATTTACATCAACAACCCACCACTAGCTTCAACA 361
DB 310 ATCAGCTTCTCAATCTGTCTGAATTTACATCAACAACCCACCACTAGCTTCAACA 369
QY 362 ACACAGAGATCAAGTCAAACTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCA 421
DB 370 ACACAGAGATCAAGTCAAACTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGTCA 429
QY 422 ACCCAACACCAAGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 481
DB 430 ACCCAACACCAAGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 489
QY 482 CCCAATATGATTTTCACTTGAAGTGTAACTTTGATCTGACGATATGACGACGAC 541
DB 490 CCCAATATGATTTTCACTTGAAGTGTAACTTTGATCTGACGATATGACGACGAC 549
QY 542 AATCAACCTGCTGGCTATCTGCAAAAGATTAACCAACCAACCAACCAACCAACCA 601
DB 550 AATCAACCTGCTGGCTATCTGCAAAAGATTAACCAACCAACCAACCAACCAACCA 609
QY 602 ACCGACAGGCTACAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAAC 661
DB 610 ACCGACAGGCTACAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAAC 669
QY 662 ACCGACAGGCTACAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAACCA 721
DB 670 ACCGACAGGCTACAAAAACCAACCTTCAAGACCAACCAACCAACCAACCAACCAAC 729
QY 722 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 781
DB 730 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 789

QY 782 ACAAGTCAAAATGGAACCTTCACATCAACCTCTCGAAGGACAAATCAAGCCCTTCTCA 841
DB 790 ACAAGTCAAAATGGAACCTTCACATCAACCTCTCGAAGGACAAATCAAGCCCTTCTCA 849
QY 842 GTCTCCACAACATCCGAGACCCATACAAACCTTATCTCCACCAACCAACCAACCAAC 901
DB 850 GTCTCCACAACATCCGAGACCCATACAAACCTTATCTCCACCAACCAACCAACCAAC 909
QY 902 TAGTATT 909
DB 910 TAGTATT 917
RESULT 11
AR080424 894 bp DNA linear PAT 31-AUG-2000
LOCUS AR080424
DEFINITION Sequence 28 from patent US 5968776.
ACCESSION AR080424
VERSION AR080424.1 GI:10007159
KEYWORDS
SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 894)
AUTHORS Klein,M.H., Du,R.-P. and Ewasysbyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL Patent: US 5968776-A 28 19-Oct-1999;
FEATURES
source Location/Qualifiers
BASE COUNT 364 a 288 c 93 g 149 t
ORIGIN
Query Match 97.2%; Score 894; DB 6; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.6e-184;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 8 ATGTCCAAAACAGAGACCAACGACGCTAGACACTAGAAAAGACCTGGACACTCTC 67
DB 1 ATGTCCAAAACAGAGACCAACGACGCTAGACACTAGAAAAGACCTGGACACTCTC 60
QY 68 AATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGTAGACAA 127
DB 61 AATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGTAGACAA 120
QY 128 ATCATATTCATTCATTCGCAATGATTAATCTCACTTATATATATAGCCATCATTA 187
DB 121 ATCATATTCATTCATTCGCAATGATTAATCTCACTTATATATATAGCCATCATTA 180
QY 188 TTCAATAGCTTCGCAACCAACCAAGTCAACACTGCAATCATATCAAGATGCAACA 247
DB 181 TTCAATAGCTTCGCAACCAACCAAGTCAACACTGCAATCATATCAAGATGCAACA 240
QY 248 ACCGACATTAACAAACCAACCCCAACATACCTGACGATCTGACGATCTGACGATCTG 307
DB 241 ACCGACATTAACAAACCAACCCCAACATACCTGACGATCTGACGATCTGACGATCTG 300
QY 308 TTCTCCAAATCTGTGAAATTTACATCAACCAACCAACCAACCAACCAACCAACCAAC 367
DB 301 TTCTCCAAATCTGTGAAATTTACATCAACCAACCAACCAACCAACCAACCAACCAAC 360
QY 368 GGAGTCAAGTCAAAACCTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 427
DB 361 GGAGTCAAGTCAAAACCTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 420
QY 428 ACACAACCCAGCAAGCCCACTTCAAAACCAACCAACCAACCAACCAACCAACCAACCA 487
DB 421 ACACAACCCAGCAAGCCCACTTCAAAACCAACCAACCAACCAACCAACCAACCAACCA 480
QY 488 AATGATTTTCACTTGTGAGTGTATTAATCTTGTACCTGACACATATGAGCAACATCCA 547

RESULT 12	AR092548	AR092548	894 bp	DNA	linear	PAT 06-SEP-2000
LOCUS	AR092548					
DEFINITION	Sequence	28	from patent US 5998169.			
ACCESSION	AR092548					
VERSION	AR092548.1		GI:10019302			
KEYWORDS	Unknown.					
SOURCE						

JOURNAL
Patent: US 5998169-A 28 07-DEC-1999;
FEATURES

source	location/Qualifiers
1. .894	
BASE COUNT	/organism="unknown"
364 a	288 c 93 g 149 t
ORIGIN	

Query Match	97.28;	Score 894;	DB 6;	Length 894;
Best Local Similarity	100.0%;	Pred. No. 1.6e-184;		
Matches 894;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

8 ATGTCCAAAACAGGACCAACGACCGCTAAGACACTAGAAAGACCTGGGACACTCTC 67

68 AATCATTTATATTATCATATCATCGGGCTATTATAAGCTTAATATGCTAAATGCTGAGAGAA 127
 1 ATGTCGCAAAAACAAAGGACCAACGACCGCTAGACACATAGAAAGAAGACCTGGACACTCTC 60

Db 61 AATCATTTATTTTCATATCATCGGCTATATAAGTTAATCTTAATCTGTACACAA 120

128 ATGACATTATTCATTCTGGCATGATTAATCTCAACTTCACTTATTAATTACAGCCATCATTA 187

188 TCCATAGCCCTGGGCAAAACGACAACATCATAAATCGTGAATTTTCGCTG
Db 121 ATCACATTATCATCTTCGSCAATGATTAATCTCAACTTCACCTTAATTAATTCAGGCCATCATA 180

181 TTCTATAGCCTCGGCAACCACAAAGTCACTAATCACTGCAATCATACAAGATGCACAA 240

248 AGCCGATCAAGACACACACCACACATACCTCACTCAGATCCTGGAATCAGC 307

Db 241 AGCCAGATCAAGAACACCAACCCACACATACCTCAGTCAGGATCCTCAGCTTGGATTCAGC 300

RESULT 13	RSNICE	LOCUS	RSNICE	8510 bp ss-RNA	linear	VRL 29-NOV-2000a
DEFINITION				Human respiratory syncytial virus nonstructural prot:in (1C)		

nonstructural protein (1B), major nucleocapsid (N), phosphoprotein (P), protein (M), 1A (1A), G (G), protein (F) and envelope-associated protein (22k) gene, complete cds. M1186 K01459 K02719 K03348 K03349 M1217 M1244 M11487 M1505

VERSION M11514 M11631 M12966
 M1486.1 GI:333925
 KEYWORDS envelope-associated protein; fusion glycoprotein; major nucleocapsid protein; major surface glycoprotein; matrix protein; nonstructural protein; phosphoprotein

[illegible]

AUTHORS	Elango, N. and Venkatesan, S.
TITLE	Amino acid sequence of human respiratory syncytial virus nucleocapsid protein
JOURNAL	Nucleic Acids Res. 11 (17), 5941-5951 (1983)
ABSTRACT	

MEDLINE	83299261
PUBMED	6310521
REFERENCE	2 (bases 3211 to 4157)
AUTHORS	Satake, M. and Venkatesan, S.
TITLE	Nucleotide sequence of the gene encoding respiratory syncytial

JOURNAL
Virus matrix protein
J. Virol. 50 (1), 92-99 (1984)
MEDLINE
84138836
PUBMED
6699948
REFERENCE
3 (bases 2288 to 3191)

AUTHORS Satake, M., Eliango, N. and Venkatesan, S.
TITLE Sequence analysis of the respiratory syncytial virus phosphoprotein gene
JOURNAL J. Virol. 52 (3), 991-994 (1984)
MEDLINE 85033973
PUBMED 6548527
REFERENCE 4 (bases 5602 to 7500)
AUTHORS Collins, P.L., Huang, Y.T. and Wertz, G.W.
TITLE Nucleotide sequence of the gene encoding the fusion (F) glycoprotein of human respiratory syncytial virus
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (24), 7683-7687 (1984)
MEDLINE 85088471
PUBMED 6096849
REFERENCE 5 (bases 7551 to 8510)
AUTHORS Collins, P.L. and Wertz, G.W.
TITLE The envelope-associated 22k protein of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polytranscript
JOURNAL J. Virol. 54 (1), 65-71 (1985)
MEDLINE 85135082
PUBMED 3838351
REFERENCE 6 (bases 8 to 1050; 7554 to 8506)
AUTHORS Eliango, N., Satake, M. and Venkatesan, S.
TITLE mRNA sequence of three respiratory syncytial virus genes encoding two nonstructural proteins and a 22k structural protein
JOURNAL J. Virol. 55 (1), 101-110 (1985)
MEDLINE 85237684
PUBMED 4009789
REFERENCE 7 (bases 5602 to 7423)
AUTHORS Eliango, N., Satake, M., Colligan, J.E., Norrby, E., Camargo, E. and Venkatesan, S.
TITLE Respiratory syncytial virus fusion glycoprotein: nucleotide sequence of mRNA, identification of cleavage activation site and amino acid sequence of N-terminus of F1 subunit
JOURNAL Nucleic Acids Res. 13 (5), 1559-1574 (1985)
MEDLINE 85215565
PUBMED 2987829
REFERENCE 8 (bases 4630 to 5543)
AUTHORS Satake, M., Colligan, J.E., Eliango, N., Norrby, E. and Venkatesan, S.
TITLE Respiratory syncytial virus envelope glycoprotein (G) has a novel structure
JOURNAL Nucleic Acids Res. 13 (21), 7795-7812 (1985)
MEDLINE 86067198
PUBMED 4069997
REFERENCE 9 (bases 4627 to 5544)
AUTHORS Wertz, G.W., Collins, P.L., Huang, Y., Gruber, C., Levine, S. and Ball, L.A.
TITLE Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82 (12), 4075-4079 (1985)
MEDLINE 85216636
PUBMED 3858865
REFERENCE 10 (bases 4173 to 4571)
AUTHORS Collins, P.L. and Wertz, G.W.
TITLE The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript
JOURNAL Virology 141 (2), 283-291 (1985)
MEDLINE 86098645
PUBMED 3879976
REFERENCE 11 (bases 1 to 528; 552 to 1050)
AUTHORS Collins, P.L. and Wertz, G.W.
TITLE Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus
JOURNAL Virology 143 (2), 442-451 (1985)
MEDLINE 86045905
PUBMED 2998021
REFERENCE 12 (bases 1081 to 2277)
AUTHORS Collins, P.L., Anderson, K., Langer, S.J. and Wertz, G.W.
TITLE Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus
JOURNAL Virology 146 (1), 69-77 (1985)
MEDLINE 85301974
PUBMED 3839952

REFERENCE 13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578)
AUTHORS Collins, P.L., Dickens, L.E., Buckler-White, A., Olmsted, R.A., Spriggs, M.K., Camargo, E. and Collingh, K.V.
TITLE Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4594-4598 (1986)
MEDLINE 86259643
PUBMED 3460060
COMMENT Reprint for [2] and clean copy sequence for [12], [5] kindly provided by P. Collins, 21-Feb-1986.
Respiratory syncytial virus is a negative-strand RNA pleomorphic enveloped virus of the genus Pneumovirus. The matrix protein has no homology with the matrix proteins of other negative-stranded RNA viruses, implying that RS virus has undergone extensive evolutionary divergence. Two unidentified reading frames which potentially encode proteins were located: one overlaps the matrix protein and the other the F protein. The biological significance of these two reading frames is not clear. The positive strand is shown.

FEATURES
source location/Qualifiers
1..8510
/organism="Human respiratory syncytial virus"
/strain="A2"
/db_xref="taxon:11250"
/lab_host="HEp-2 cell"
1..528
/gene="1C"
1..528
/gene="1C"
55..474
/gene="1C"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAB59850.1"
/db_xref="GI:333926"
/translation="MGNSLSMTKVRQLONFQNDDEVAL KTCYTDKLIHLNALAKA VHTIKGLVFEVHVTSSDICRNNTIVKSNFTTPYVQNGCTIEMMELTHCSQPN GLDDNCEIKFSKRLSDSTMTYNWOLSELGDLNP"
171
/gene="1C"
/note="c in [11]; t in [6]"
/replace="t"
552..1050
/gene="1B"
552..1050
/gene="1B"
584..958
/gene="1B"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAB59851.1"
/db_xref="GI:333927"
/translation="MDTTHNDNFQRLMTDMPLSLETITSLTRDITTHKEFTYLN HECLVRKLDEQATFEFLVNTKEMKLHKVGSYKVKYETVRYGTFPMPIFNHGF LEGLGIRPKTKHTPIYKYDLNP"
747
/gene="1B"
/note="a in [11]; g in [6]"
/replace="g"
769
/gene="1B"
/note="g in [11]; a in [6]"
/replace="a"
1081..2277
/gene="N"
1081..2277
/gene="N"
1096..2271
/gene="N"
/codon_start=1
/product="major nucleocapsid"

variation
variation
variation
gene
mRNA
CDS
variation
gene
mRNA
CDS

```
/protein_id="AAB59852.1"
/db_xref="GI:333928"
/translation="MALSKYKLMNDITLNKQDLSSTKYTORSTGDSITDPNPDVQKHI
NRKCGMLLITEDNHRFTGLIGLTMASRIKGRBDITKILRDACYHVKAGVDVTHRO
DINGEMKEEVLTLATFTEIQTINIEESKSYKMKJEMGEVAPEDRSPCCGMI I
LCTALAVITKLAGDSRGLAVIRANVNLKMKRYKGLLEPDILANSFYEEKEMHI
FIDVPEVFGIAOSTRGSGRVEGIFAGLPMNAATGACQVMLRMGVLAKEVKNIMLGHS
VOAEMQVVEYEAOKLGEAGFYHILNPKRSLISTLOPHERSSVLCGNAAGLGIIM
GEGRTPRNDLIDAKAKAYAEOLKENGIVINSVLDTAELELEIKHQLNPKDNDVEL"
2289..3196
/gene="p"
/gene="p"
2289..3196
/gene="p"
2302..3027
/gene="p"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB59853.1"
/db_xref="GI:333929"
/translation="MEKFAPEFHGEDNARRATKFILESIKGFSPRPDKKDDSTISVN
SIDIEVTKESPITSNSTIINPTNETDGTAGNKNYORKPLVSEKEDPPSPDPSKLY
KETIEDNNNEESSYSYEINDOTNDNTARLDREKLSFICMLHTLVYASGPT
SARDGIRDMIGLRREMIEKIRTEALMTNDRLEAARLNRESEKMAKDTSDVSLNP
TSEKLNMLEGSDSDNLSLEDP"
```

Query Match 92.8%; Score 853.4; DB 14; Length 8510;
Best Local Similarity 95.5%; Pred. No. 9.6e-176;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
OY 1 TGCAAACATGTCCAAAACAGACCAAGCACCCTAGACACTGTGAAAAGACCTGGGA 60
Db 4634 TGCAAACATGTCCAAAACAGACCAAGCACCCTAGACACTGTGAAAAGACCTGGGA 4693
OY 61 CACTCTCAATCATTTATTTATTCATCATCGGGCTTATATTAATTAATCTTAATCTGT 120
Db 4694 CACTCTCAATCATTTATTTATTCATCATCGGGCTTATATTAATTAATCTTAATCTGT 4753
OY 121 AGCAACAATCATTTATTCATTTGCGAATGATATCTCAACTCTTATATTAATTAACG 180
Db 4754 AGCAACAATCATTTATTCATTTGCGAATGATATCTCAACTCTTATATTAATTAACG 4813
OY 181 CATCATATTCATAGCTCGGCAACCAAGTCACATTAACCAATGATCATCAAGA 240
Db 4814 CATCATATTCATAGCTCGGCAACCAAGTCACATTAACCAATGATCATCAAGA 4873
OY 241 TGCACAGCAGCATCAAGACCAACCCCAATACCTCACTCAAGATCCTCAGCTTG 300
Db 4874 TGCACAGCAGCATCAAGACCAACCCCAATACCTCACTCAAGATCCTCAGCTTG 4933
OY 301 AATCAGCTTCTCAATCTGTCTGAATTTACATCAACAACCCACCATCTAGCTTCAAC 360
Db 4934 AATCAGCTTCTCAATCTGTCTGAATTTACATCAACAACCCACCATCTAGCTTCAAC 4993
OY 361 AACACGAGAGTCAAGTCAAGCTGCAACCGCAACAGTCAGAGCTTAATAACCAACAAC 420
Db 4994 AACACGAGAGTCAAGTCAAGCTGCAACCGCAACAGTCAGAGCTTAATAACCAACAAC 5053
OY 421 AACCCAAACACAAACCAGACGCCACTGACAAACAGCCCAACCAACCAACAACA 480
Db 5054 AACCCAAACACAAACCAGACGCCACTGACAAACAGCCCAACCAACCAACAACA 5113
OY 481 ACCCATATATGATTTTCACTTGGAAGTGTTTAACTTTGACCTGACAGCATATGACGAA 540
Db 5114 ACCCATATATGATTTTCACTTGGAAGTGTTTAACTTTGACCTGACAGCATATGACGAA 5173
OY 541 CAATCCAAACCTGCGGCTATCTGCAAAAAGAAATACCAAAAACCAAGAAACAAAC 600
Db 5174 CAATCCAAACCTGCGGCTATCTGCAAAAAGAAATACCAAAAACCAAGAAACAAAC 5233
OY 601 CACCAACAGGCTACAAAAAACCAACTCTCAAGCAACCAAAAAAGATCTCAAACTCA 660
Db 5234 CACTACCAAGGCTACAAAAAACCAACTCTCAAGCAACCAAAAAAGATCTCAAACTCA 5293
```

```
OY 661 AACCACTAAACCAAGGAAGTACCCACCAACCAAGCCACAGAGAAGACCAACCATCAAC 720
Db 5294 AACCACTAAATCAAAAGGAAGTACCCACCAACCAAGCCACAGAGAAGACCAACCATCAAC 5353
OY 721 CACCAAAACCAAGCATACAGTACACTGCTGACCAACCAACCAACCAAGCAATCCAAACT 780
Db 5354 CACCAAAACCAAGCATATATTAACCTACTACCTGCTGACCAACCAACCAAGCAATCCAAACT 5413
OY 781 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
Db 5414 CACAAGTCAAAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 5473
OY 841 AGCTCCCAACATACCGACACCCATACACACCTTCATCTGACCAACCAACCAAGCA 900
Db 5474 AGCTCCCAACATACCGACATCCATACCAACCTTATCTCCACCAACCAACCAAGCA 5533
OY 901 GTAGTTATTTAAAAA 919
Db 5534 GTAGTTACTTAAAAACAT 5552
```

RESULT 14

AR093219 15222 bp DNA lineat PAT 08-SEP-2000
LOCUS AR093219
DEFINITION Sequence 23 from patent US 5998602.
ACCESSION AR093219
VERSION AR093219.1 GI:10019970
KEYWORDS

SOURCE
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 15222)
AUTHORS Torrence,P.F., Silverman,R.Hugh., Cirino,N.Mario., Li,G. and Xiao,W.

TITLE Rnase L activators and antisense oligonucleotides effective to treat RSV infections
JOURNAL Patent: US 5998602-A 23 07-DEC-1999;
FEATURES location/Qualifiers
source 1..15222

BASE COUNT 5924 a 2704 c 2356 g 4238 t
ORIGIN

Query Match 92.8%; Score 853.4; DB 6; Length 15222;
Best Local Similarity 95.5%; Pred. No. 9.2e-176;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
OY 1 TGCAAACATGTCCAAAACAGACCAAGCACCCTAGACACTGTGAAAAGACCTGGGA 60
Db 4681 TGCAAACATGTCCAAAACAGACCAAGCACCCTAGACACTGTGAAAAGACCTGGGA 4740
OY 61 CACTCTCAATCATTTATTTATTCATCATCGGGCTTATATTAATTAATCTGT 120
Db 4741 CACTCTCAATCATTTATTTATTCATCATCGGGCTTATATTAATTAATCTGT 4800
OY 121 AGCAACAATCATTTATTCATCATCGGGCTTATATTAATTAATCTGT 180
Db 4801 AGCAACAATCATTTATTCATCATCGGGCTTATATTAATTAATCTGT 4860
OY 181 CATCATATTCATAGCTTCGCAACCAAGTCACACTAACAATCTCAATATACAGA 240
Db 4861 CATCATATTCATAGCTTCGCAACCAAGTCACACTAACAATCTCAATATACAGA 4920
OY 241 TGCACAAAGCAGATCAAGAACCAACCCCAATACCTGACGATGAGATCTCAGCTTG 300
Db 4921 TGCACAAAGCAGATCAAGAACCAACCCCAATACCTGACGATGAGATCTCAGCTTG 4980
OY 301 AATCAGCTTCTCAATCTGTCTGAATTTACATCAACAACCAACCAACCAATCTAGCTTCAAC 360
Db 4981 AATCAGCTTCTCAATCTGTCTGAATTTACATCAACAACCAACCAACCAATCTAGCTTCAAC 5040
OY 361 AACACCAAGATCAAGTCAACCAAGCACCAGTCAAGATCAAGATCAAGATCAAGATCAAG 420
```



```

Db      5041 AACACGAGGAGTCAATCAACCCCTGCATCCAAACAGTCAACGCCAACCAACAAACAAC 5100
QY      421 AACCAAACAAACCCAGCAGAGCCCACTCAAAAACACGCCCAAAACCAACCAACCA 480
Db      5101 AACTCAAAACAAACCCAGCAGAGCCCAACCAAAACCAACCCCAACCAACCAACCA 5160
QY      481 ACCCAATTAATGATTTTCACTGCAAGTGTTTTAACTTTTACCCCTGAGCATATGACGAA 540
Db      5161 ACCCAATTAATGATTTTCACTGCAAGTGTTTTAACTTTTACCCCTGAGCATATGACGAA 5220
QY      541 CAATCCAAACCTGCTGGGCTATCTGCAAAAAGAAATATCAAAAACAAAACAGAAAGAAAC 600
Db      5221 CAATCCAAACCTGCTGGGCTATCTGCAAAAAGAAATATCAAAAACAAAACAGAAAGAAAC 5280
QY      601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGAACCAACCAAAAAAAGATCTCAACCTCA 660
Db      5281 CACTACCAACCCCAACAAAAAACCAACCCCTCAAGAACCAACCAAAAAAAGATCTCAACCTCA 5340
QY      661 AACCACTAAACCAAGGAAGTACCCACACCAAGCCCAAGAGAGCAACCAATCAACAC 720
Db      5341 AACCACTAAATCAAGGAAGTACCCACACCAAGCCCAAGAGAGCAACCAATCAACAC 5400
QY      721 CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGAAATCAAAACT 780
Db      5401 CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGAAATCAAAACT 5460
QY      781 CACAAGCAATGAAGAAACCTTCACCTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840
Db      5461 CACAAGCAATGAAGAAACCTTCACCTCAACCTCTCGAAGGCAATCTAAGCCCTTCTCA 5520
QY      841 AGTCTCCACAACATCCGAGCAGCCATCAACACCTGATCTCCACCAACCAACCAACGCGCA 900
Db      5521 AGCTCTCAACAACATCCGAGCAGCCATCAACACCTGATCTCCACCAACCAACCAACGCGCA 5580
QY      901 GTAGTTATTAAAAAAA 919
Db      5581 GTAGTTACTTAAAAACATA 5599

RESULT 15
HRU50362      15222 bp      RNA      linear      VRL-20-NOV-1997
LOCUS      Human respiratory syncytial virus, mutant cp-RSV, complete genome.
DEFINITION      U50362
ACCESSION      U50362.1 GI:2627296
VERSION      1
KEYWORDS      Human respiratory syncytial virus.
SOURCE      Human respiratory syncytial virus.
ORGANISM      Paramyxoviridae; Pneumovirinae; Pneumovirus.
REFERENCE      1 (bases 1 to 15222)
AUTHORS      Connors, M., Crowe, J.E. Jr., Firestone, C.Y., Murphy, B.R. and
              Collins, P.L.
TITLE      A cold-passaged, attenuated strain of human respiratory syncytial
              virus contains mutations in the F and L genes
JOURNAL      Virology 208 (2), 478-484 (1995)
MEDLINE      95266253
PUBMED      7747420
REFERENCE      2 (bases 1 to 15222)
AUTHORS      Crowe, J.E. Jr., Firestone, C.Y., Whitehead, S.S., Collins, P.L. and
              Murphy, B.R.
TITLE      Acquisition of the ts phenotype by a chemically mutagenized
              cold-passaged human respiratory syncytial virus vaccine candidate
              results from the acquisition of a single mutation in the polymerase
              (L) gene
JOURNAL      Virus Genes 13 (3), 269-273 (1996)
MEDLINE      97187925
PUBMED      9035372
REFERENCE      3 (bases 1 to 15222)
AUTHORS      Whitehead, S.S.
TITLE      Direct Submission
JOURNAL      Submitted (29-FEB-1996) RVS, LID, NIAID, Bldg 7, Rm 118, 7 Center
              Drive, MSC 0720, Bethesda, MD 20892, USA

```

```

FEATURES
Source
location/Qualifiers
1..15222
/organism="Human respiratory syncytial virus"
/strain="A2"
/db_xref="taxon:11250"
/note="mutant cp-RSV"
45..576
/gene="NS1 (1C)"
45..576
/gene="NS1 (1C)"
/product="nonstructural protein NS1"
99..518
/gene="NS1 (1C)"
/codon_start=1
/product="nonstructural protein NS1"
/db_xref="GI:2627297"
/translation="MGSNSLSMTKYRLQNLFDNDVALLKICNYDNLHLTNALAKA
VHRTKNGIVFVHVTSSDICPNNTIVKSNFTTPVLQNGYIWMNELTRCSOPN
GLDDNCEIKFSKRLSDSTMTNMYNOLSEILGFDLNP"
596..1098
/gene="NS2 (1B)"
596..1098
/gene="NS2 (1B)"
/product="nonstructural protein NS2"
628..1002
/gene="NS2 (1B)"
/codon_start=1
/product="nonstructural protein NS2"
/db_xref="GI:2627298"
/translation="MDTTHNDNPORLMTDMRPLSEITITSLRDIITHKRTIYLN
HECTVRLDERQATFPLVNEKMLKHKVSTKYKTEITNTYGTGPPMDFINHDGF
LECGIRKPTHTPIYKYDLNP"
1125..2327
/gene="N"
1125..2327
/gene="N"
/product="major nucleocapsid"
1140..2315
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/db_xref="GI:2627299"
/translation="MALSKYKNDLTUKDQLSSKTYTIOSTGSDTDPNVYOKHI
NKLGLMLITEDANHKFTGLIGLIYASRSKREDTIKLIDAGIHKRANKVDYTTBQ
DINKEKMFVLTLASLTLEIOLNIELESRSYKKMLKEGEVAPERYHDSPOGMITI
LCIALVITKILAADRSGLTAVIRANNVLRNKKRYKGLPLDIOLANSFEVEKEKHP
FIDVFVHFGIAGOSTRGSRVEGIFAGLFNNAYGAGVLRMGVLAKSININMIGHS
VOAMEOVVEVYEYAKKLGEGAGFYHLNPKASLSLTQEPHSSVVLGNAGLIGM
GEYRGTPRNDPLYNAKAKYAKQLEKENVINVSVDLTAELIAIKHOLNPKDNDVEL"
2329..3242
/gene="P"
2329..3242
/gene="P"
/product="phosphoprotein"
2346..3071
/gene="P"
/codon_start=1
/product="phosphoprotein"
/db_xref="GI:2627300"
/translation="MEKFAPEFHGEDANNRATKFLSIKGFSPDKPKKKSIIYSN
SIDLEVTRESPTNSITINFTNETDTAGKPKPYQKRPVSEKEDPTSDNPSKLY
KETIETEDNNEESSYSYEETINDOTNDITAKDRIDEKUSEIIGHIHTLVASAGPT
SARDGIRDAWAGLREEMLEKIRTLALMTNRLLEAMRLNESEKMAKDTSDVSLNP
TSEKLNLLGENDSDNDLSEDF"
3252..4209
/gene="M"
3252..4209
/gene="M"
/product="Matrix protein M"

```



```

Db 5341 AACCACTAAATCAAAAGGAGTACCCACCAAGCCACAGAAAGGCCAATCAACAC 5400
QY 721 CACCAAAACAAACATCACTACACTGCTCACCACACACAGAAATCCAAAAT 780
Db 5401 CACCAAAACAAACATCACTACACTGCTCACCACACACAGAAATCCAAAAT 5460
QY 781 CACAAGTCAAATGGAACCTTCACATCACTCCTCCGAAAGGCAATCTAAGCCCTCTCA 840
Db 5461 CACAAGTCAAATGGAACCTTCACATCACTCCTCCGAAAGGCAATCTAAGCCCTCTCA 5520
QY 841 AGTCTCCACAAATCCGAGACCCGATCAACCTCATCTCCACCCAAACACAGSCCA 900
Db 5521 AGTCTCCACAAATCCGAGACCCGATCAACCTCATCTCCACCCAAACACAGSCCA 5580
QY 901 GTAGTTATTAAAAAAA 919
Db 5581 GTAGTTACTTAAAAACATA 5599

```

Search completed: May 1, 2003, 11:59:35
 Job time : 1875.32 secs

us-09-462-816-1.rng

OM nucleic - nucleic search, using sw model

11542.374 Million cell updates/sec

```
Sequence:      920  
1 tgcacaacatgtcccaaac.....gtagtlatlaaaaaaaaaa 920
```

Gapop 10.0 , Gapext 1.0

number of hits satisfying chosen parameters: 4370478

```
maxLlull DB seq length: 200000000000
```

Maximum Match 1008
Listing first 45 summaries

1	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1980.DAT *
2	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1981.DAT *
3	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1982.DAT *
4	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1983.DAT *
5	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1984.DAT *
6	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1985.DAT *
7	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1986.DAT *
8	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1987.DAT *
9	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1988.DAT *
10	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1989.DAT *
11	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1990.DAT *
12	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1991.DAT *
13	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1992.DAT *
14	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1993.DAT *
15	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1994.DAT *
16	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1995.DAT *
17	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1996.DAT *
18	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1997.DAT *
19	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1998.DAT *
20	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1998.DAT *
21	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA1999.DAT *
22	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA2000.DAT *
23	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA2001.DAT *
24	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA2002.DAT *
25	/SIDS2/gcgdata/geneseq/geneseqn_emb1/NA2002.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description
	1	97.0	100.0	920	14	AA045686	Respiratory syncyt
	2	92.0	100.0	920	20	AA080421	Membrane bound G p
	3	85.7	93.2	935	8	AA070784	Sequence encoding
	4	85.7	93.2	935	8	AA070784	HRSV glycoprotein
	5	85.6	93.0	935	13	AA029623	HRSV glycoprotein
	6	85.3	92.8	15222	18	AA018440	Human respiratory
	7	85.3	92.8	15223	19	AA063430	Respiratory syncyt
	8	85.4	92.8	15223	18	AA017551	Respiratory syncyt
	9	85.3	92.8	15223	21	AA088743	Respiratory syncyt

10	848.8	92.2	91.8	22	AAC88494	Human RSV G-proc
11	837.8	92.1	897	24	AA820145	Respiratory syncy
12	830.4	90.3	15210	20	AAK59703	Polynucleotide seq
13	715	77.7	715	20	AAK08422	G protein gene fr
14	637.4	69.3	696	19	AAV188298	Respiratory syncy
15	441.4	48.0	15225	19	AAV17552	Respiratory syncy
16	441.4	48.0	15225	21	AAAB88744	Human respiratory
17	441.4	48.0	15225	24	AAAD5520	RSV isolate 18537
18	422.2	45.9	15229	19	AAAB18276	Polynucleotide sequ
19	422.2	45.9	15229	20	AAAZ22910	DNA encoding the l
20	422.2	45.9	15229	20	AAK35268	RSV isolate 2B wil
21	411.4	44.7	15218	19	AAV18275	Nucleotide sequen
22	411.4	44.7	15218	20	AAAZ22909	Nucleotide sequen
23	411.4	44.7	15218	20	AAAZ22914	DNA encoding the l
24	411.4	44.7	15218	20	AAK35267	Nucleotide sequen
25	411.4	44.7	15219	19	AAV18277	DNA encoding the l
26	411.4	44.7	15219	19	AAV18278	RSV vaccine 2B33F
27	411.4	44.7	15219	19	AAV18279	RSV vaccine 2B30L
28	411.4	44.7	15219	19	AAV18280	RSV revertant 2B33
29	411.4	44.7	15219	20	AAAZ22911	RSV revertant 2B22
30	411.4	44.7	15219	20	AAAZ22912	Nucleotide sequen
31	411.4	44.7	15219	20	AAAZ22913	Nucleotide sequen
32	411.4	44.7	15219	20	AAK35269	Nucleotide sequen
33	411.4	44.7	15219	20	AAK35270	DNA encoding the l
34	411.4	44.7	15219	20	AAK35271	DNA encoding the l
35	411.4	44.7	15219	20	AAK35272	DNA encoding the l
36	212.8	23.1	1050	17	AAAT31617	Phocetes Streptococ
37	212.8	23.1	1050	22	AAAB47167	Nucleotide sequen
38	212.8	23.1	1356	22	AAAB80153	Nucleotide sequen
39	210.8	22.9	303	16	AAAT03486	RSV subgroup A clo
40	210.8	22.9	303	17	AAAT27066	RSV subgroup A wile
41	210.8	22.9	303	17	AAAT27073	RSV subgroup A wile
42	210.8	22.9	303	17	AAAT31649	RSV subgroup A wile
43	210.8	22.9	303	21	AAAZ30478	NSP ratory Syncyt
44	210.8	22.9	303	21	AAAZ59882	RSV protein anti
45	210.8	22.9	303	22	AAAT78459	DNA encoding a G2N
						Nuc isolate sequen

ALIGNMENTS

RESULT 1	
AA045686	
ID	AA045686 standard; DNA; 920 BP.
XX	
AC	AA045686;
XX	
DT	13-JAN-1994 (first entry)
XX	
DE	Respiratory syncytial virus (RSV) G gene.
XX	
KW	P1V; RSV; multimeric; hybrid; pathogen; chimeric p
XX	
XX	ds.
OS	Respiratory syncytial virus.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	8..901
FT	/*tag= a
FT	/product= RSV G protein
FT	89..176
FT	/*tag= b
FT	/label= Transmembrane anchor domain
XX	
PN	MO9314207-A.
XX	
PD	22-JUL-1993.
XX	
PF	05-JAN-1993;
XX	
PR	93WO-CA00001.
XX	
PR	06-JAN-1992;
XX	
XX	92GB-0000117.
PA	(CONN-) CONNAUGHT LAB LTD.

```

XX      Ewasysbyn ME, Klein MH;
PI      WPI: 1993-243222/30.
DR      P-PSDB: AAR39286.
XX      Multimeric hybrid genes and their chimeric proteins - are
PT      vaccines against multiple pathogenic infections e.g.
PT      para-influenza virus and respiratory syncytial virus
XX
PS      Claim 11: Figure 7A-7D; 80pp; English.
XX
CC      A novel multimeric hybrid gene is used as a vaccine. The gene
CC      consists of two gene sequences which are linked and encode antigenic
CC      regions, these two sequences being derived from two different
CC      pathogens (parainfluenza virus (PIV) and respiratory syncytial virus
CC      (RSV)). The gene sequences that are particularly used are those
CC      which encode PIV-3 F and HN proteins (AA045683, AA045684) and RSV F and
CC      G proteins (AA045685, AA045686).
XX
SQ      Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;
Query Match      100.0%; Score 920; DB 14; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.2e-211;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY      1 TGCACAAATGTCACAAACAGACACGCGCTAAGACACTAGAAAAGACCTGCGA 60
DB      1 TGCACAAATGTCACAAACAGACACGCGCTAAGACACTAGAAAAGACCTGCGA 60
OY      61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAAGTTAAATCTTAACTCTGT 120
DB      61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAAGTTAAATCTTAACTCTGT 120
OY      121 AGCACAATACATATATTCATCTGCGAATGATATTCATCACTTATATTAATACAGC 180
DB      121 AGCACAATACATATATTCATCTGCGAATGATATTCATCACTTATATTAATACAGC 180
OY      121 AGCACAATACATATATTCATCTGCGAATGATATTCATCACTTATATTAATACAGC 180
DB      121 AGCACAATACATATATTCATCTGCGAATGATATTCATCACTTATATTAATACAGC 180
OY      181 CATCATATTCATATGCTCGGCAACGACAAAGTCACTACACTGACATCATACAGA 240
DB      181 CATCATATTCATATGCTCGGCAACGACAAAGTCACTACACTGACATCATACAGA 240
OY      181 CATCATATTCATATGCTCGGCAACGACAAAGTCACTACACTGACATCATACAGA 240
DB      181 CATCATATTCATATGCTCGGCAACGACAAAGTCACTACACTGACATCATACAGA 240
OY      241 TGCACACAGCGATTAAGACACAAACCCACATACCTGACACTGAGATCTGAGCTGG 300
DB      241 TGCACACAGCGATTAAGACACAAACCCACATACCTGACACTGAGATCTGAGCTGG 300
OY      241 TGCACACAGCGATTAAGACACAAACCCACATACCTGACACTGAGATCTGAGCTGG 300
DB      241 TGCACACAGCGATTAAGACACAAACCCACATACCTGACACTGAGATCTGAGCTGG 300
OY      301 AATCAGCTCTGCAATCTGTGTAATTCATACAAACACACACATCTAGCTTCAAC 360
DB      301 AATCAGCTCTGCAATCTGTGTAATTCATACAAACACACACATCTAGCTTCAAC 360
OY      301 AATCAGCTCTGCAATCTGTGTAATTCATACAAACACACACATCTAGCTTCAAC 360
DB      301 AATCAGCTCTGCAATCTGTGTAATTCATACAAACACACACATCTAGCTTCAAC 360
OY      361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACACAAAC 420
DB      361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGACTTAAACACACAAAC 420
OY      421 AACCCAAACGACACGACGAGCCACTACAAACGCAAAAGCAAAACCAACCAACAA 480
DB      421 AACCCAAACGACACGAGCCACTACAAACGCAAAAGCAAAACCAACCAACAA 480
OY      481 ACCCAATATGATTTTCACTTCAAGTGTGTTAACTTTGTAACCTGACGATATGACGAA 540
DB      481 ACCCAATATGATTTTCACTTCAAGTGTGTTAACTTTGTAACCTGACGATATGACGAA 540
OY      541 CAATTCACCTGTGTGGCTATCTGCAAAAGATATCAACAAACCAAGGAAAGAAAC 600
DB      541 CAATTCACCTGTGTGGCTATCTGCAAAAGATATCAACAAACCAAGGAAAGAAAC 600
OY      601 CACACACAGCTGACAAACCAACCTTCAAGACACACCAACCAACCAACCAACCTCA 660
DB      601 CACACACAGCTGACAAACCAACCTTCAAGACACACCAACCAACCAACCAACCTCA 660
OY      661 AACCACTAAACCAAGAGAGTACCAACCAAGCCACAGAGAGCCCAACCAACCAACCA 720
DB      661 AACCACTAAACCAAGAGAGTACCAACCAAGCCACAGAGAGCCCAACCAACCAACCA 720

```

```

OY      721 CACCAAAACAAATCATCACTACATGCTGCACCAACACACAGGAAATCCAAACT 780
DB      721 CACCAAAACAAATCATCACTACATGCTGCACCAACACACAGGAAATCCAAACT 780
OY      781 CACAAGTCAAAATGAGAAACCTTCACACTCAACCTGCTCCGAGGAATCTTAGCCCTTCTCA 840
DB      781 CACAAGTCAAAATGAGAAACCTTCACACTCAACCTGCTCCGAGGAATCTTAGCCCTTCTCA 840
OY      841 AGCTCTCAACATATCGACACCCATCAACACCTCATCTCCAGCCACACACAGGCA 900
DB      841 AGCTCTCAACATATCGACACCCATCAACACCTCATCTCCAGCCACACACAGGCA 900
OY      901 GTAGTTATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
DB      901 GTAGTTATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920

```

RESULT 2
AA08421
AA08421 standard; cDNA; 920 BP.

AX08421;

28-JUN-1999 (first entry)

Membrane bound G protein gene of respiratory syncytial virus.

G protein: respiratory syncytial virus; RSV; recombinant vector;

vaccine; immune response; immunogenicity; tPA; antibody;

tissue plasminogen activator; ss.

Respiratory syncytial virus (RSV).

Key Location/Qualifiers

CDS 8..904 /*tag- a /product- "Membrane bound G protein"

W09904010-A1.

28-JAN-1999.

16-JUL-1998; 98WO-CA00697.

18-JUL-1997; 97US-0896442.

(CONN-) CONNAUGHT LAB LTD.

Klein MH, LI X, Sambhara S;

WPI: 1999-132254/11.

P-PSDB: AAM96313.

Immunogenic composition for generating antibodies against

respiratory syncytial virus - comprises non-replicating vector

containing the protein G sequence, useful in protective vaccines and

to raise antibodies for diagnosis

Claim 3; Figure 2; 67pp; English.

The respiratory syncytial virus (RSV) G protein can be used in

vaccines by inserting the G protein gene into a non-replicating

vector. The G protein is placed under the control of alternative

signal and expression sequences, for example the chimeric G protein

activator (tPA). The recombinant vector may also comprise sequences

upstream of the G protein gene which enhance the G protein's

immunoprotective ability. The resulting immunogenic composition will

generate antibodies directed against the RSV G protein when

administered to a host organism. The composition is useful as a

vaccine to immunise against RSV-associated disease, particularly

resulting in a balanced Th1/Th2 immune response and for raising Ab,

CC by usual immunisation and cell fusion methods.

XX Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;

Query Match 100.0%; Score 920; DB 20; Length 920;

Best Local Similarity 100.0%; Pred. No. 3.2e-211;

Matches 920: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TCGAACAATGTCTCAAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACCTGGGA 60
DB 1 TCGAACAATGTCTCAAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACCTGGGA 60
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
DB 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
QY 121 AGCAAAATTCATATTCATTTCTGGCAATGATTAATCTCAACTTCACTTATTAATAGAGC 180
DB 121 AGCAAAATTCATATTCATTTCTGGCAATGATTAATCTCAACTTCACTTATTAATAGAGC 180
QY 181 CATCATTTTCATAGCTCGGCAACGACCAAGTCAACTACATCACTGCAATCATACAGA 240
DB 181 CATCATTTTCATAGCTCGGCAACGACCAAGTCAACTACATCACTGCAATCATACAGA 240
QY 241 TGCACAAGCCAGATTCAGAACACACACCCCAACATACCTCAGATCCTCAGCTTGG 300
DB 241 TGCACAAGCCAGATTCAGAACACACACCCCAACATACCTCAGATCCTCAGCTTGG 300
QY 301 AATCAGCTTCTCAGATCTGCTGTAATTCATCATCAAAACGACCACTACTAGCTTGG 360
DB 301 AATCAGCTTCTCAGATCTGCTGTAATTCATCATCAAAACGACCACTACTAGCTTGG 360
QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
DB 361 AACACGAGAGTCAAGTCAAACTGCAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
QY 421 AACCAAAACGACAGCCGAGAAAGCCACTACAAAACGACCAAGCAACCAACCAACCAAC 480
DB 421 AACCAAAACGACAGCCGAGAAAGCCACTACAAAACGACCAAGCAACCAACCAACCAAC 480
QY 481 ACCCAATATGATTTTCACTTCAAGTGTAACTTTGACCCGCGAGCATATGACAGCA 540
DB 481 ACCCAATATGATTTTCACTTCAAGTGTAACTTTGACCCGCGAGCATATGACAGCA 540
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATATCAAAACCAAAACCAAGAAAGAAAC 600
DB 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATATCAAAACCAAAACCAAGAAAGAAAC 600
QY 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACACCAACCAAAAGATCTCAACCTCA 660
DB 601 CACCACCAAGCTTACAAAACCAACCTTCAAGACACCAACCAAAAGATCTCAACCTCA 660
QY 661 AACCACTAAACCAAGGAGTACCCACCAAGCCACAGAGAGCCCAACCTCAACAG 720
DB 661 AACCACTAAACCAAGGAGTACCCACCAAGCCACAGAGAGCCCAACCTCAACAG 720
QY 721 CACCAAAACCAATCAGACACTACACTGCTCAGCAACCAACAGAGAAATTCAAAAC 780
DB 721 CACCAAAACCAATCAGACACTACACTGCTCAGCAACCAACAGAGAAATTCAAAAC 780
QY 781 CACCAAGTCAATGGAACCTTCACTCAACCTTCCGAAGGCAATCTAAGCCCTTCTCA 840
DB 781 CACCAAGTCAATGGAACCTTCACTCAACCTTCCGAAGGCAATCTAAGCCCTTCTCA 840
QY 841 AGCTCTCACAACATCCGAGCAACCCATACAAACCTATCTCAGCCCAACCAAGAGCGCA 900
DB 841 AGCTCTCACAACATCCGAGCAACCCATACAAACCTATCTCAGCCCAACCAAGAGCGCA 900
QY 901 GTAGTTATTTAAAAA 920
DB 901 GTAGTTATTTAAAAA 920

```

```

RESULT 3
ID AAN70784
XX AAN70784 standard; cDNA; 935 BP.
AC AAN70784;
XX 05-APR-1991 (first entry)
DT
XX
XX Sequence encoding human respiratory syncytial virus (HRSV) A2 strain
DE G protein.
XX
XX Vaccine; ss.
XX
XX Human respiratory syncytial virus (HRSV).
FH Key Location/Qualifiers
FT CDS 16..913
FT /tag= a
XX
PN W08704185-A.
PN
XX 16-JUL-1987.
PD
XX
XX 23-DEC-1986; 86MO-US02756.
PR
XX 14-JAN-1986; 86US-0818740.
XX
XX (UYNC-) UNIV OF N CAROLINA.
XX (WERT/) WERTZ G W.
XX (WERT/) WERTZ G W.
XX
XX WPI: 1987-206300/29.
XX P-PSDB; AAF70845.
XX
XX Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.
XX
XX Disclosure; Chart 13; 57pp; English.
XX
XX A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women, or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.
XX
XX
SQ Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match 93.2%; Score 857.6; DB 8; Length 935;
Best Local Similarity 95.8%; Pred. No. 3e-196;
Matches 881: Conservative 0; Mismatches 39; Indels 0; Gaps 0;
QY 1 TCGAACAATGTCTCAAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACCTGGGA 60
DB 9 TCGAACAATGTCTCAAAACAGGACCAAGCCGCTTAAGACACTAGAGAAAGACCTGGGA 68
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 120
DB 69 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAATCTGT 128
QY 121 AGCAAAATTCATATTCATTTCTGGCAATGATTAATCTCAACTTCACTTATTAATAGAGC 180
DB 129 AGCAAAATTCATATTCATTTCTGGCAATGATTAATCTCAACTTCACTTATTAATAGAGC 188
QY 181 CATCATTTTCATAGCTCGGCAACGACCAAGTCAACTTAAGCACTCAATCATACAGA 240
DB 189 CATCATTTTCATAGCTCGGCAACGACCAAGTCAACTTAAGCACTCAATCATACAGA 248
QY 241 TGCACAAGCCAGATTCAGAACACACCCCAACATACCTCAGAGTCTCAGCTTGG 300
DB 249 TGCACAAGCCAGATTCAGAACACACCCCAACATACCTCAGAGTCTCAGCTTGG 308
QY 301 AATCAGCTTCTCAGATCTGCTGTAATTCATCAACCAAGCAACCACTACTAGCTTCAAC 360

```

```

Db      309  AATCAGTCCCTTAATCGCTGAAATTACATCAAAATTCACCATCTAGCTTCAAC 368
OY      361  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 420
Db      369  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 428
OY      421  AACCCACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAACAA 480
Db      429  AACCTCAACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAAC 488
OY      481  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 540
Db      489  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 548
OY      541  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 600
Db      549  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 608
OY      601  CACACACAGCCTACAAAAAACCAACCTTCAAGAACCAACCAAAAGATCTCAAACTCA 660
Db      609  CACTACCAAGCCCAAAAAAACCAACCTTCAAGAACCAACCAAAAGATCTCAAACTCA 668
OY      661  AACCACTAAACCAAGAGTAACCCACCAAGCCCAAGCCCAAGAGCCCAACCATCAAC 720
Db      669  AACCACTAAATCAAGAGTAACCCACCAAGCCCAAGCCCAAGAGCCCAACCATCAAC 728
OY      721  CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGAAATCCAAACT 780
Db      729  CACCAAAACCAACATCAACATCACTGCTCAACCAACCAACCAAGAAATCCAAACT 788
OY      781  CACAAGTCAATGGAACCTTCACTCAACCTCTCCGAAGCAATCTAAGCCCTTCTCA 840
Db      789  CACAAGTCAATGGAACCTTCACTCAACCTCTCCGAAGCAATCTAAGCCCTTCTCA 848
OY      841  AGTCTCCACAAATCCGAGCAGCCATCAACCCCTCATCTCCAAACCAACCAAGCCCA 900
Db      849  AGTCTCTACAAATCCGAGTCCGATCAACCACTTCACTCCAAACCAACCAAGCCCA 908
OY      901  GTAGTTATTTAAAAAAA 920
Db      909  GTAGTTATTTAAAAAAA 928

RESULT 4
AAV18736
ID      AAV18736 standard; cDNA; 935 BP.
XX      AC      AAV18736;
XX      DT      11-JUN-1998 (first entry)
XX      DE      HRSV glycoprotein G cDNA.
XX      KM      HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine; ss.
XX      OS      Human respiratory syncytial virus.
XX      FH      Key
XX      FT      CDS      16..912
XX      FT      /tag= a
XX      FT      /product= glycoprotein_G
XX      PN      US5716823-A.
XX      PD      10-FEB-1998.
XX      PF      12-MAY-1997; 97US-0854783.
XX      PR      13-JUL-1988; 8805-0218737.
XX      PR      14-JAN-1986; 8605-0818740.
XX      PR      23-DEC-1986; 86WO-US02756.
XX      PR      11-JUN-1992; 92US-0897171.

```

```

PR      12-MAY-1997; 97US-0854783.
XX      PA      (PHAA ) PHARMACIA & UPJOHN CO.
XX      PI      Collins PL, Wertz GW;
XX      DR      WPI; 1998-144802/13.
XX      PT      P-PSDB; AAM47605.
XX      PS      Production of human respiratory syncytial virus glyco-protein F or G
XX      CC      - by culturing eukaryotic host cells transfected with corresponding
XX      CC      DNA
XX      CC      Example 1; Columns 27-28; 17pp; English.
XX      CC      The present sequence was used in the development of a novel method
XX      CC      for the production of human respiratory syncytial virus (HRSV)
XX      CC      glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
XX      CC      culturing eukaryotic host cells transfected with an isolated DNA
XX      CC      sequence encoding HRSV gpF or gpG. The gp can be used to prepare
XX      CC      vaccines against HRSV.
SO      Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match      93.28; Score 857.6; DB 19; Length 935;
Best Local Similarity 95.88; Pred. No. 3e-196;
Matches 881; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

OY      1  TGCAACATGTCGCAAAAAACAGGACCAACGACCGCTAAGACACTAGAAAAAGACTGGGA 60
Db      9  TGCAACATGTCGCAAAAAACAGGACCAACGACCGCTAAGACACTAGAAAAAGACTGGGA 68
OY      61  CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATATGTTAATCTTAATCTGT 120
Db      69  CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATATATGTTAATCTTAATCTGT 128
OY      121  AGCAAAATTCACATTATTCATTTGCGCAATGATATATCAACTCACTTAAATTAACAGC 180
Db      129  AGCAAAATTCACATTATTCATTTGCGCAATGATATATCAACTCACTTAAATTAACAGC 188
OY      181  CATCATATTCATAGCGCTGGGCAACCAACAAAGTCACACTAACAAGTCAATCAACAGA 240
Db      189  CATCATATTCATAGCGCTGGGCAACCAACAAAGTCAACCAACCAACTCAATCATACAGA 248
OY      241  TGCAACAAAGCAGATCAAGAACACACACCCCAACTACTCTCATGAGATCCCTGAGCTGG 300
Db      249  TGCAACAAAGCAGATCAAGAACACACACCCCAACTACTCTCATGAGATCCCTGAGCTGG 308
OY      301  AATCAGCTTCTCCAAATCTGTCTGAAATATACACAAACCAACCAACTACTAGCTTCAAC 360
Db      309  AATCAGTCCCTCTAATCCGTCTGAAATATACACAAATCAACCAACTACTAGCTTCAAC 368
OY      361  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 420
Db      369  AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGATAAATAACACAAAC 428
OY      421  AACCCACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAACAA 480
Db      429  AACCTCAACACACACCCAGCAGCCCACTACAAACAGCCCAAAACCAACCAACCAAC 488
OY      481  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 540
Db      489  ACCCAATATGATTTCACTTGAAGTGTCTTAACTTTGACCTTCAGCATATGACGAA 548
OY      541  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 600
Db      549  CAATCCAACTCTGCGGTATCTGCAAAAGAAATACCAACCAAAACCAAGAAAGAAC 608
OY      601  CACACACAGCCTACAAAAAACCAACCTTCAAGAACCAACCAAAAGATCTCAAACTCA 660
Db      609  CACTACCAAGCCCAAAAAAACCAACCTTCAAGAACCAACCAAAAGATCTCAAACTCA 668
OY      661  AACCACTAAACCAAGAGTAACCCACCAAGCCCAAGAGCCCAACCATTAACAC 720

```

```

|||||
Db 669 AACCACTAAATCAAGAGAGTACCCACCAAGCCACAGAGAGCAACATCAACAC 728
OY 721 CACCAAAACAAATCAGCACTACACGTCGACCAACACACAGGAATCCAAACT 780
Db 729 CACCAAAACAAATCAGCACTACACGTCGACCAACACAGGAATCCAAACT 788
OY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTCTCA 840
Db 789 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTCTCA 848
OY 841 AGTCTCACAACATCCGAGCACCATCACAACCCCTATCTCCACCCACACACAGCCCA 900
Db 849 AGTCTCACAACATCCGAGCACCATCACAACCCCTATCTCCACCCACACAGCCCA 908
OY 901 GTAGTTATTAATAAAAAA 920
Db 909 GTAGTTACTTAATAAAAAA 928

RESULT 5
AAO29623
ID AAO29623 standard; DNA; 935 BP.
XX
AC AAO29623;
XX
DT 03-MAR-1993 (first entry)
XX
DE HSRV glycoprotein G (gpG).
XX
KM Vaccine: human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
KM major capsid protein; N: ss.
XX
OS Human respiratory syncytial virus strain A2.
XX
FH Key Location/Qualifiers
FT CDS 16..912
FT /tag= a
FT /label= G-protein
FT misc_feature 16..22
FT /tag= b
FT /note= "Oligonucleotide used to probe for
FT misc_feature 898..912 full length CDNA"
FT /tag= C
FT /note= "Oligonucleotide used to specifically
FT prime the reverse transcription reaction
FT for making the first strand of the CDNA"
XX
PN US5149650-A.
XX
PD 22-SEP-1992.
XX
PF 14-JAN-1986; 86US-0818740.
XX
PR 14-JAN-1986; 86US-0818740.
PR 13-JUL-1988; 88US-0218737.
XX
PA (UYNC-) UNIV NORTH CAROLINA.
XX
PI Collins PL; Wertz GW;
XX
DR MPI; 1992-340247/41.
DR P-PSDB; AAR25302.
XX
PT Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments
XX
PS Disclosure; Page 18; 21pp; English.
XX
CC The sequences of mRNA encoding HRSV structural proteins are given in
CC AAO29622-26. The proteins are F, G, 22K, 9.5K and major capsid

```

```

CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.
XX
SQ Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;

Query Match
Best Local Similarity 93.0%; Score 856; DB 13; Length 935;
Matches 880; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

OY 1 TGCAAACATGTCCAAAACAAAGACCAAGCGCCCTAAGACACTAGAAAAGCCTGGGA 60
Db 9 TGCAAAACATGTCCAAAACAAAGACCAAGCGCCCTAAGACACTAGAAAAGCCTGGGA 68
OY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTGT 120
Db 69 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTGT 128
OY 121 AGCACAATACATATATCATTCATTCGCAATGATATCTCACTTATATATACAGC 180
Db 129 AGCACAATACATATATTCATTCGCAATGATATCTCACTTATATATATGAGC 188
OY 181 CATCATATTCATATAGCCTCGGCAACCAAGTCACTAACAATCGCATCATACAGA 240
Db 189 CATCATATTCATATAGCCTCGGCAACCAAGTCACTAACAATCGCATCATACAGA 248
OY 241 TGCACAAGCCGATCAAGACACACACCCCAATACCTCACTAGATCTCTACCTTGG 300
Db 249 TGCACAAGCCGATCAAGACACACACCCCAATACCTCACTAGATCTCTACCTTGG 308
OY 301 AATCAGCTTCTCCANTCTGTCTGANAATTACATCAACAACCCACCATATACAGCTCAAC 360
Db 309 AATCAGTCCCTCTATATTCGCTGGAATTTACATCAACAATCAACCATATACAGCTCAAC 368
OY 361 AACACGAGAGTCAAGTCAAACTGCAACCCACAACTCAAGCTAAATTAACAC 420
Db 369 AACACGAGGCGTCAAGTCAAACTGCAACCCACAACTCAAGCTAAATTAACAC 428
OY 421 AACCCAAACACACCCGACCAAGCCGACATCAAAACAGCCAAACACCAACCA 480
Db 429 AACTCAAAACACACCCGACCAAGCCGACCAAAACAGCCAAACACCAACCA 488
OY 481 ACCCAATATGATTTTCACTTGAAGTGTTTTAACTTGTATACCTGACGATATGACCAA 540
Db 489 ACCCAATATGATTTTCACTTGAAGTGTTTTAACTTGTATACCTGACGATATGACCAA 548
OY 541 CAATGCAACCTGCTGGGTATCTGCAAAAGATACCAACAAACAAACCAAGAAAGAAAC 600
Db 549 CAATGCAACCTGCTGGGTATCTGCAAAAGATACCAACAAACAAACCAAGAAAGAAAC 608
OY 601 CACCACCAAGCCTTACAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAACTCA 660
Db 609 CACTACCAAGCCCAACAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAACTCA 668
OY 661 AACCACTAAACCAAGAAAGTACCCACACCAAGCCCAAGAAAGCCCAACATCAACAC 720
Db 669 AACCACTAAATCAAGAAAGTACCCACACCAAGCCCAAGAAAGCCCAACATCAACAC 728
OY 721 CACCAAAACAAATCACTACATGCTGACCAACCAACCAACCAAGAAATCCAAACT 780
Db 729 CACCAAAACAAATCACTACATGCTGACCAACCAACCAACCAAGAAATCCAAACT 788
OY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTCTCA 840
Db 789 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTCTCA 848
OY 841 AGTCTCACAACATCCGAGCACCATCACAACCTTATCTCCACCCACACAGCCCA 900
Db 849 AGTCTCACAACATCCGAGCACCATCACAACCTTATCTCCACCCACACAGCCCA 908
OY 901 GTAGTTATTAATAAAAAA 920
Db 909 GTAGTTACTTAATAAAAAA 928

```


XX Collins PL;
 PT WPI: 1997-212893/19.
 XX
 PT Infectious respiratory syncytial virus particles - useful for
 treatment of RSV or gene therapy of upper respiratory tract diseases
 XX
 PS Claim 46; Page 43-51; 66pp; English.
 CC
 CC A human respiratory syncytial virus (RSV) anti-genome sequence
 CC (AAT63430) is the 5' to 3' positive-sense sequence of RSV; the genome
 CC itself is negative-sense. It was synthesised in segments by RT-PCR
 CC using intracellular RSV mRNA or genomic RNA isolated from purified
 CC viruses as template. Restriction site markers were introduced by
 CC incorporating the changes into the primers used for RT-PCR. The
 CC recombinant sequence can be expressed with a nucleocapsid protein,
 CC a nucleocapsid phosphoprotein, a large polymerase protein and an
 CC RNA elongation factor to produce isolated infectious RSV particles
 CC or antigenome can also be used as a vector for gene therapy of the
 CC upper respiratory tract.
 CC
 XX
 XX Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other:
 Query Match 92.8%; Score 853.4; DB 18; Length 15223;
 Best Local Similarity 95.5%; Pred. No. 6.9e-195;
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCACATGTCGCAAAACAGGACGACGACCGCTTAAGACCTTGAAAGACCTGGGA 60
 DB 4682 TGCACATGTCGCAAAACAGGACGACGACCGCTTAAGACCTTGAAAGACCTGGGA 4741
 QY 61 CACTCTCAATCATTTATATATCATATCATGCGGCTTATTAAGTTAAATCTTAATCTCT 120
 DB 4742 CACTCTCAATCATTTATATCATATCATGCGGCTTATTAAGTTAAATCTTAATCTCT 4801
 QY 121 AGCACAATTCACATTTATCATTTCTGGCATGATAATCTCAACTCTTAATTAATGACG 180
 DB 4802 AGCACAATTCACATTTATCATTTCTGGCATGATAATCTCAACTCTTAATTAATGACG 4861
 QY 181 CATCATATTCATAGCTCGGCAAAACAGGACGACGACCGCTTAAGACCTTGAAAGACCTGGGA 240
 DB 4862 CATCATATTCATAGCTCGGCAAAACAGGACGACGACCGCTTAAGACCTTGAAAGACCTGGGA 4921
 QY 241 TGCACACGCGAGATCAAGACACACCCCAATACCTACTCAGATCCTCAGCTTGG 300
 DB 4922 TGCACACGCGAGATCAAGACACACCCCAATACCTACTCAGATCCTCAGCTTGG 4981
 QY 301 AATCAGCTTCTCCATCTCTGTAATTTACATCACAACCCACCATCTACTAGCTTCAAC 360
 DB 4982 AATCAGCTTCTCCATCTCTGTAATTTACATCACAACCCACCATCTACTAGCTTCAAC 5041
 QY 361 AACACGAGGAGTCAAGTCAACCTGCAACGACGACGACGACGACGACGACGACGACGACGAC 420
 DB 5042 AACACGAGGAGTCAAGTCAACCTGCAACGACGACGACGACGACGACGACGACGACGACGAC 5101
 QY 421 AACCCAAACACACCCAGAGCCCACTACAAACAGCCCAAAACAAACCAACCAACAA 480
 DB 5102 AACCTCAAAACACACCCAGAGCCCACTACAAACAGCCCAAAACAAACCAACCAACAA 5161
 QY 481 ACCCAATATGATTTTCACTTGGAAAGTGTTAACCTTGGACCTGCAGCATATGACGCAA 540
 DB 5162 ACCCAATATGATTTTCACTTGGAAAGTGTTAACCTTGGACCTGCAGCATATGACGCAA 5221
 QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAAGATACCAAAACCAAGGAAAGAAAC 600
 DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAAGATACCAAAACCAAGGAAAGAAAC 5281
 QY 601 CACGACCAAGGCTACAAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAACTCA 660
 DB 5282 CACTACCAAGGCTACAAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAAACTCA 5341

QY 661 AACCACTAAACCAAGGAGTACCCACACCAAGCCCAAGAGACCAACCATC AACAC 720
 DB 5342 AACCACTAAATCAAGGAGTACCCACACCAAGCCCAAGAGACCAACCATC AACAC 5401
 QY 721 CACCAAAACCAATCAACATACACTGCTGCAACCAACCAACCAAGGAAATTC:AAACT 780
 DB 5402 CACCAAAACCAATCAACATACACTGCTGCAACCAACCAACCAAGGAAATTC:AGAACT 5461
 QY 781 CACAGTCAAAATGGAACCTTCCACTCAACCTCTCCCAAGCAATCTAAC TTTCTCA 840
 DB 5462 CACAGTCAAAATGGAACCTTCCACTCAACCTCTCCCAAGCAATCTAAC TTTCTCA 5521
 QY 841 ACTCTCACAACATCCGAGACCCATCAACCCCTCATCTCCACCAACCAACCAAGGCA 900
 DB 5522 ACTCTCTCAACATCTCCGAGTACCCATCAACCCCTCATCTCCACCAACCAACCAAGGCA 5581
 QY 901 GTAGTTATTTAAAAAAA 919
 DB 5582 GTAGTTACTTTAAAAACATA 5600

RESULT 8
 AAV17553
 ID AAV17553 standard; cDNA; 15223 BP.
 XX
 XX AAV17553;
 AC
 XX
 XX 20-JUL-1998 (first entry)
 DT
 XX
 XX
 DE Respiratory syncytial virus antigenome.
 XX
 XX RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.
 KW
 XX Human respiratory syncytial virus p46.
 OS
 XX WO9802530-A1.
 PN
 XX
 PD 22-JAN-1998.
 XX
 PF 15-JUL-1997; 97WO-US12269.
 XX
 PR 23-MAY-1997; 97US-0047634.
 PR 15-JUL-1996; 96US-0021773.
 PR 09-MAY-1997; 97US-0046141.
 XX
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Bukreyev AA, Collins PL, Juhász K, Murphy BR, Teng MN;
 DR Whitehead SS;
 XX
 XX WPI: 1998-110579/10.
 PT
 PT Attenuated respiratory syncytial virus vaccines - useful to protect
 PT individuals against RSV infection
 XX
 PS Example 7; Page 188-195; 238pp; English.
 XX
 XX This is the 5'-3' positive sequence nucleotide sequence of
 XX respiratory syncytial virus (RSV) p46. The genome is
 XX negative-sense; the complete nucleotide sequence of the
 XX wild-type B-1 virus has also been determined (see AAV17552).
 XX A novel infectious recombinant RSV comprises a RSV genome or
 XX antigenome, a major nucleocapsid (N) protein, a nucleocapsid
 XX phosphoprotein (P), a large polymerase protein (L), and a RNA
 XX polymerase elongation factor, where the recombinant RSV has at
 XX least two attenuating mutations; one of the mutations specifying a
 XX temperature-sensitive (ts) substitution at amino acid Phe521,
 XX Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts
 XX nucleotide substitution in the gene-start sequence of gene M2.
 XX Also claimed are: (1) an isolated infectious RSV particle which
 XX comprises a recombinant RSV (antigenome, N, P, and L proteins, a
 XX RNA polymerase elongation factor, where the (anti)genome is modified:
 XX (1) to ablate or modulate expression of a SH, NS1, NS2 or G gene or

CC a cis-acting regulatory sequence; and (ii) by a termination codon
 CC introduced within a selected gene, or by a change in sequence,
 CC position or presence of a GS or GE transcription signal relative to
 CC the selected gene; (2) an expression vector; and (3) an RSV strain
 CC selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR
 CC 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),
 CC cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or
 CC B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated
 CC attenuated recombinant RSV and RSV particles are used in a vaccine
 CC to stimulate the immune system of an individual to induce
 CC protection against RSV. The expression vector of (2) is used for
 CC the production of infectious attenuated RSV particles.

XX
 XX
 XX Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other;

Query Match 92.8%; Score 853.4; DB 19; Length 15223;
 Best Local Similarity 95.5%; Pred. No. 6,9e-195;
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TGCAGCATGTCGCAAAAAGAGGACGACGCGCTAGACCTGAGAAAGACCTGGGA 60
 DB 4682 TGCAGCATGTCGCAAAAAGAGGACGACGCGCTAGACCTGAGAAAGACCTGGGA 4741
 QY 61 CACTGCAATCATTTATTTATTCATATCATGCGGCTTATATAGTTAAATCTGT 120
 DB 4742 CACTGCAATCATTTATTTATTCATATCATGCGGCTTATATAGTTAAATCTGT 4801
 QY 121 AGCACAATACATTTATTCATTTCTGGCAATGATTAATCTCACTTATTAATTACAGC 180
 DB 4802 AGCACAATACATTTATTCATTTCTGGCAATGATTAATCTCACTTATTAATTACAGC 4861
 QY 181 CATCATTTATGAGCCCTGGCAACCCAGCAAGTCACACTACATGCAATCATACAG 240
 DB 4862 CATCATTTATGAGCCCTGGCAACCCAGCAAGTCACACTACATGCAATCATACAG 4921
 QY 241 TGCACAAGCCAGATCAAGACCAACCCGACATACCTTCACTAGAGTCTGAGCTTGG 300
 DB 4922 TGCACAAGCCAGATCAAGACCAACCCGACATACCTTCACTAGAGTCTGAGCTTGG 4981
 QY 301 AATCAGTCTGCAATGCTGTGTAATTAATCATGCAAAACGACACATATAGCTTCAAC 360
 DB 4982 AATCAGTCTGCAATGCTGTGTAATTAATCATGCAAAACGACACATATAGCTTCAAC 5041
 QY 361 AACACGAGAGTCAAGTCAAAACCTGCAACCCGACATGCAAGTCAAGTCAAGTCAAG 420
 DB 5042 AACACGAGAGTCAAGTCAAAACCTGCAACCCGACATGCAAGTCAAGTCAAGTCAAG 5101
 QY 421 AACCCAACACACCCGACGCAAGCCCACTACAAAAACGCAAAACCAACCAACCAAC 480
 DB 5102 AACCCAACACACCCGACGCAAGCCCACTACAAAAACGCAAAACCAACCAACCAAC 5161
 QY 481 ACCCAATTAAGATTTTACCTGCAAGTGTGTTAACTTTGTACCTGCGAGCTATGACGAA 540
 DB 5162 ACCCAATTAAGATTTTACCTGCAAGTGTGTTAACTTTGTACCTGCGAGCTATGACGAA 5221
 QY 541 CAATCAACCTGCTGGGCTATGTCGAAAAGAAATACCAAAACCAAGGAAAGAAAC 600
 DB 5222 CAATCAACCTGCTGGGCTATGTCGAAAAGAAATACCAAAACCAAGGAAAGAAAC 5281
 QY 601 CACGACCAAGCTTCAAAAAACCAACCTTCAAGACAAACCAAAAAAATCTCAAACTCA 660
 DB 5282 CACGACCAAGCTTCAAAAAACCAACCTTCAAGACAAACCAAAAAAATCTCAAACTCA 5341
 QY 661 AACCACTAAACCAAGGAGTACCCACCAAGCCGACAGAGGAGCAACCTCAACAC 720
 DB 5342 AACCACTAAACCAAGGAGTACCCACCAAGCCGACAGAGGAGCAACCTCAACAC 5401
 QY 721 CACCAAAACCAACATCAACATCACTGCTGACCAACACACACAGAAATCCAAACT 780
 DB 5402 CACCAAAACCAACATCAACATCACTGCTGACCAACACACAGAAATCCAAACT 5461
 QY 781 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCTCGAAGGCAATTAAGCCCTTCTCA 840

DB 5462 CACAAGTCAAATGGAACCTTCCACTCAACCTCTCTCGAAGGCAATCAACGACCTTCTCA 5521
 QY 841 AGTCTCAACATCCGACGACCCATCAACACCTTATCTTCACCCAAACACACGCGCA 900
 DB 5522 AGTCTCTCAACATCCGAGTACCATCAACCTTCACTTCACCCAAACACACGCGCA 5581
 QY 901 GTAGTTATTAATAAAAAA 919
 DB 5582 GTAGTTACTTAATAAACATA 5600

RESULT 9 AAA88743

ID AAA88743 standard; cDNA; 15223 BP.

AC AAA88743;

DT 19-FEB-2001 (first entry)

DE Respiratory syncytial virus D46 5'-3' positive sense sequence.

XX RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.

OS Chimeric - Human respiratory syncytial virus.

OS Chimeric - Bacteriophage T7.

XX Synthetic.

PH Key Location/Qualifiers

FT variation replace(4,6)

FT /*tag- a

FT mutation 1099

FT /*tag- b

FT /*note= "single C insertion to create A1111 site in the NS2-N intergenic region"

FT mutation replace(1139,A)

FT /*tag- c

FT /*note= "creates NcoI site in N gene nontranslated region"

FT mutation replace(1140,G)

FT /*tag- d

FT /*note= "creates NcoI site in N gene nontranslated region"

FT mutation replace(5612,A)

FT /*tag- e

FT /*note= "creates StuI site in G/F intergenic region"

FT mutation replace(5616,A)

FT /*tag- f

FT /*note= "creates StuI site in G/F intergenic region"

FT mutation replace(7560,A)

FT /*tag- g

FT /*note= "creates SphI site in F/M2 intergenic region"

PD WO20061611-A2.

XX 19-OCT-2000.

XX 31-MAR-2000; 2000WO-US08802.

XX 13-APR-1999; 99US-0291894.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Collins PL, Murphy BR, Whitehead SS;

XX WPI; 2000-679462/66.

XX Infectious chimeric respiratory syncytial virus (RSV) produced from

XX cloned nucleotide sequences, useful as a vaccine against diseases

XX caused by the virus, such as pneumonia and bronchiolitis -

XX Example 7; Page 262-268; 280pp; English.

XX The present sequence is that of the 5' to 3' positive-sense

Query Match	92.88	Score 853.4	DB 21	Length 15223
Best Local Similarly	95.58	Pred. No. 6.9e+195		
Matches 878	Conservative	0	Mismatches 41	Indels 0
				Gaps 0

OY		661	AACCACTAAACCAGGAGAATGTACCCACACCGAACGCCACAGGAAGGCACAACCATTAACAC	720
Db		5342	AACCACTAAATCAAAGGAATAATTACCCACACCAAGCCCACAGAAGAGCCAATCAACAC	5401
OY		721	CACCAAAAACAATCACAACACTACACGTTCACCAACACACACACAGCAATCCAAACT	780
Db		5402	CACCAAAAAACAATCATCTAATTAATCTACTACTTCACTCTCCACACACACAGAAATCTAGAACT	5461
OY		781	CACAAGTCAAATGGAAAACCTTGCCACTCACACCTCCTCCGAAAGGCAATCTAAGCCCCTTCTCA	840
Db		5462	CACAAGTCAAATGGAAAACCTTGCCACTCAACTTGCTCCGAAAGGCAATCCAAAGCCCTTCTCA	5521
OY		841	AGTCTGCACACACATCCGAGCACCACATCAACAACCTCATCTCTCCACCCACAACACAGCCCA	900
Db		5522	AGCTCTTAGCAACATCCGAGATCCATCAACCTTCATCTCCACCAACACACAGCCCA	5581
OY		901	GTAGTTATTATTTTTAAAAAAA 919 	
Db		5582	GTAGTTACTTTAAAAACATGA 5600	
	RESULT 10			
ID	AAC8494			
XZ	AAC88494 standard; DNA; 918 BP.			
XX	AAC88494:			
DT	12-MAR-2001 (first entry)			
XX	Human RSV G-protein gene.			
DE	Chimeric; respiratory syncytial virus; RSV; immunize; js.			
KM	Unidentified.			
OS	WO200068392-A1.			
XX	16-NOV--2000.			
PN	10-MAY-2000; 2000MO-US12582.			
XX	PR			
XX	11-MAY-1999; 99US-O133536.			
PA	(UNIT) UNITV ILLINOIS POUND.			
XX	Buetow DE, Korban SS, Sandhu J, Krasnyanski SF;			
PI	WPI: 2001-122707/13.			
DR				
PT	Chimeric nucleic acid construct for immunizing animals and humans			
XX	- against respiratory syncytial virus (RSV), comprises a sequence adapted			
PS	for expression in plants and a RSV protein or peptide coding sequence			
XX				
XX	Disclosure: Fig 10; 67pp; English.			
CC	The present invention relates to a chimeric nucleic acid construct			
CC	comprising: a nucleotide sequence adapted for protein expression in			
CC	plants; and a respiratory syncytial virus (RSV) coding sequence			
CC	encoding an RSV protein or an antigenic protein or peptide of RSV.			
CC	The construct can be used to immunize animals and humans against			
CC	respiratory syncytial virus. The use of transgenic plants to			
CC	generate the antigen allows the production of greater amounts of			
CC	antigen.			
SQ	Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;			
	Query Match 92.2%; Score 848.2; DB 22; Length 918; Best local Similarity 95.8% ; Pred. No. 5.4e-194; Matches 871; Conservative 0; Mismatches 38; Indels 0; Gaps 0;			
OY	1 TCSCAACATGTCTCAAAAACAGGAGCAACGCGCTAGAGACTAGAAAAGACCTMGGGA 60			


```

Db 61 AATCATTTATATATCATATCATGTCGTATATTAAGTTAAATCTTAATCTGTAGACAAA 120
QY 128 ATCATATATATCATATTTGGCAATGATATATCAATTCACCTTATATATATACAGCATATA 187
Db 121 ATCATATATATCATATTTGGCAATGATATATCAATTCACCTTATATATATATGAGCCATATA 180
QY 188 TTCAATGCTTCGGCAAAACCAAAAGTCACACTTAACAATCAATCATATACAGATGACACA 247
Db 181 TTCAATGCTTCGGCAAAACCAAAAGTCACACTTAACAATCAATCATATACAGATGACACA 240
QY 248 AGCCAGATCAAGAACACAAACCCCAACATACCTCACTAGATGCTCAAGTTGGAAATCAGC 307
Db 241 AGCCAGATCAAGAACACAAACCCCAACATACCTCACTAGATGCTCAAGTTGGAAATCAGC 300
QY 308 TTCTCATATCTGTCTGAATTTATATCATCAACACCAACCATACCTAGTGTTCATCAACACCA 367
Db 301 CCTCTATATCGCTGTGAATTTATATCATCAACACCAACCATACCTAGTGTTCATCAACACCA 360
QY 368 GGAATCAAGTCAAACTGCAACCCCAACAGTCAAGACTAAAAACACACACACACCA 427
Db 361 GGAATCAAGTCAAACTGCAACCCCAACAGTCAAGACTAAAAACACACACACCA 420
QY 428 ACACAAACCCCAAGCCCACTTACAAAACACAGCCCAAAACCAACCAACCAACCAAT 487
Db 421 ACACAAACCCCAAGCCCACTTACAAAACACAGCCCAAAACCAACCAACCAACCAAT 480
QY 488 AATGATTTTCACTTTCGAAGTGTTAAGTTTGTACCTGACGATATGACACACCAATGCA 547
Db 481 AATGATTTTCACTTTCGAAGTGTTAAGTTTGTACCTGACGATATGACACCAATGCA 540
QY 548 ACCTGCTGGGCTATCTGCAAAAGATTCACAAACAAAAACCAAGAAAGAAACCAACCA 607
Db 541 ACCTGCTGGGCTATCTGCAAAAGATTCACAAACAAAAACCAAGAAAGAAACCAACCA 600
QY 608 AAGCTCAAAAAAACCAACCTTCAAGACACCAACCAAAAAAGATCTCAAACTCAAACT 667
Db 601 AAGCTCAAAAAAACCAACCTTCAAGACACCAACCAAAAAAGATCTCAAACTCAAACT 660
QY 668 AAACCAAGAGAGTACCCACACCAAGCCCAAGAGAGGCAACCATCAACACCAACAA 727
Db 661 AAATCAAGAGAGTACCCACACCAAGCCCAAGAGAGGCAACCATCAACACCAACAA 720
QY 728 ACACAAATCACTACACTGCTCACCACACACACACACACAGAAATCCAAATCTCAGAGT 787
Db 721 ACACAAATCACTACACTGCTCACCACACACACACACACAGAAATCCAAATCTCAGAGT 780
QY 788 CAATATGAAACCTTTCACATCAACCTCTCGGAAGGCAATCTTAAGCCCTTCTCAAGTCTCC 847
Db 781 CAATATGAAACCTTTCACATCAACCTCTCGGAAGGCAATCTTAAGCCCTTCTCAAGTCTCT 840
QY 848 ACACATTCGAGACCCATCAACACCTCATCTCCACCAACACACACACAGCCAGTAG 904
Db 841 ACACATTCGAGACCCATCAACACCTCATCTCCACCAACACACACACAGCCAGTAG 897

```

RESULT 12
AA59703/c
ID AA59703 standard; DNA; 15210 BP.

```

XX AAX59703;
XX
XX 26-JUL-1999 (first entry)
XX
XX Polynucleotide sequence of RSV strain A2.
XX
XX Antisense oligonucleotide; negative-strand RNA virus; activator; RNase L;
XX respiratory syncytial virus; RSV; influenza; mumps; rabies; ss.
XX Respiratory syncytial virus.
XX
XX W09922742-A1.
XX
XX 14-MAY-1999.
XX
PD

```

```

XX
XX 02-NOV-1998; 98WO-US23391.
XX
XX Best Local Similarity 94.6%; Pred. No. 2.3e-169;
XX PR 03-NOV-1997; 97US-0962690.
XX
XX (CLEV-) CLEVELAND CLINIC FOUND.
XX PA (USSH ) US NAT INST OF HEALTH.
XX
XX Clinio NM, LI G, Player MR, Silverman RH, Torrence P;
XX Xiao W;
XX WPI; 1999-326917/27.
XX
XX PT New composition useful for inhibiting or treating infections against
XX negative-strand RNA virus
XX
XX PS Disclosure; Fig 1; 98pp; English.
XX
XX The specification describes a composition comprising a polynucleotide
XX consisting of an antisense oligonucleotide containing a hydroxy group,
XX complementary to the genomic or antigenomic strand of a negative-strand
XX RNA virus; and an activator of RNase L. The polynucleotide is used to
XX inhibit, or treat, infection by negative-strand RNA viruses, specifically
XX respiratory syncytial virus (RSV) but also (para)influenza, mumps, and
XX rabies. The polynucleotide can cross cell membranes without requiring
XX carriers or permeabilizing agents, and can selectively cleave the RNA
XX targeted by the oligonucleotide. The present sequence represents
XX the polynucleotide sequence of RSV strain A2.
XX
XX SQ Sequence 15210 BP; 4232 A; 2351 C; 2700 G; 5919 T; 8 other;
XX
XX Query Match 90.3%; Score 830.4; DB 20; Length 15210;
XX Best Local Similarity 94.6%; Pred. No. 2.3e-169;
XX Matches 869; Conservative 3; Mismatches 41; Indels 6; Gaps 1;
XX
QY 1 TSCAAATATGTCAAAACCAAGAGCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60
Db 10530 TSCAAATATGTCAAAACCAAGAGCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 10471
QY 61 CACTGTCAATCACTTTATATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTG 120
Db 10470 CACTGTCAATCACTTTATATTCATATTCATTCG-----THYMGTTAAATCTTAAATCTG 10417
QY 121 AGCACAATACATATTCATATTCATGCAATGATATATCTCAACTTCACTTAATATACAGC 180
Db 10416 AGCACAATACATATTCATATTCATGCAATGATATATCTCAACTTCACTTAATATACAGC 10357
QY 181 CATCATATTCATAGCTTCGGCAACACCAAGTACACACTAAACCTGCAATCATACAGA 240
Db 10356 CATCATATTCATAGCTTCGGCAACACCAAGTACACACTAAACCTGCAATCATACAGA 10297
QY 241 TSCAACAAGCCAGATCAAGAACACAAACCCCAAGATCTCATAGATCTCTCAAGTCTGG 300
Db 10296 TSCAACAAGCCAGATCAAGAACACAAACCCCAAGATCTCATAGATCTCTCAAGTCTGG 10237
QY 301 AATCAGCTTCTCCATCTGTCTGAATTTATATCATCAAAACCAACCATATCAGTTCAGC 360
Db 10236 AATCAGCTTCTCCATCTGTGAATTTATATTCATCAAAATTCACCAACCATATCAGTTCAGC 10177
QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 420
Db 10176 AACACGAGAGTCAAGTCAAACTGCAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAG 10117
QY 421 AACCAACACACACACACACAGCCCAAGCAAAACAGCCCAAAACAGCAACCAACCA 480
Db 10116 AACCAACACACACACACAGCCCAAGCAAAACAGCCCAAAACAGCAACCAACCA 10057
QY 481 ACCCAATATATATTTCACTTGAAGTGTTAAGTTTGAATCTGACCTGACGATATGACAGCA 540
Db 10056 ACCCAATATATATTTCACTTGAAGTGTTAAGTTTGAATCTGACCTGACGATATGACAGCA 9997
QY 541 CAATTCACCTGCTGGGCTATCTGCAAAAGATTCACAAACAAAAACCAAGAAAGAAAC 600

```

```
Db 9996 CAAATCCACCTGCTGGGCTATCTGCAAAAGAAATACCAAAACAAAACAGAGAAAGAAAC 9937
QY 601 CACCACCAAGCTTACAAAAAACCACTTCAAGACACCAAAAAAGATCTCAAACTCA 660
Db 9936 CACTACCAACCCCAAAAAAACCAACCTCAAGACACCAAAAAAGATCCCAAACTCA 9877
QY 661 AACCACTAAACCAAGAAAGTACCCACCAACCAAGCCACAGAGAGCCCAACATCAACAC 720
Db 9876 AACCACTAAATCAAGAAAGTACCCACCAACCAAGCCACAGAGAGCCCAACATCAACAC 9817
QY 721 CACCAAAACCAACATCAACCACTACTGCTCAACCAACCAAGAGAAATCAAAACT 780
Db 9816 CACCAAAACCAACATCACTACTACTACTACTACTACTACTACTACTACTACTACTACT 9757
QY 781 CACAAGCAAAATGAAACCTTCACTCAACCTTCCGAAAGCAATCTAAGCCCTTCTCA 840
Db 9756 CACAAGCAAAATGAAACCTTCACTCAACCTTCCGAAAGCAATCTAAGCCCTTCTCA 9697
QY 841 AGTCTCCACAACATCCGACGACCCATCAACACCTCTCAACCCAGACACACAGCCCA 900
Db 9696 AGTCTCCACAACATCCGAGTACCCATCAACACCTCTCAACCCAGACACACAGCCCA 9637
QY 901 GTAGTTATTAAAAAAA 919
Db 9636 GTAGTTACTTAAAAACATA 9618
```

RESULT 13

AAK08422 ID AAK08422 standard; cDNA; 715 BP.

AAK08422; XX

28-JUN-1999 (first entry) XX

G protein gene fragment of respiratory syncytial virus. DE

G protein; respiratory syncytial virus; RSV; recombinant vector; KM

vaccine; immune response; immunogenicity; tPA; antibody; KM

tissue plasminogen activator; ss. XX

Respiratory syncytial virus (RSV). OS

Key Location/Qualifiers XX

CDS 1..702 /tag= a XX

/product= "Secreted G protein" XX

W09904010-A1. XX

28-JAN-1999. XX

16-JUL-1998; 98WO-CA00697. XX

18-JUL-1997; 97US-0896442. XX

(CONN-) CONNAUGHT LAB LTD. XX

Klein MH, Li X, Sambhara S; XX

WPI; 1999-132254/11. XX

P-PSDB; AAW96314. XX

Immunogenic composition for generating antibodies against XX

respiratory syncytial virus - comprises non-replicating vector XX

containing the protein G sequence, useful in protective vaccines and XX

to raise antibodies for diagnosis XX

Claim 8; Figure 3; 67pp; English. XX

The respiratory syncytial virus (RSV) G protein can be used in XX

vaccines by inserting the G protein gene into a non-replicating XX

vector. The G protein is placed under the control of alternative CC

CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods.

SQ Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;

Query Match 77.7%; Score 715; DB 20; Length 715;

Best Local Similarity 100.0%; Pred. No. 4.6e-162;

Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 CACAAGTCACTACATCACTGCAATCTATACAGATGCAACACCCAGATCAAGACACA 265

Db 1 CACAAGTCACTACATCACTGCAATCTATACAGATGCAACACCCAGATCAAGACACA 60

QY 266 ACCCAAGATACCTGACTCAGGATCTCAGCTTGGATTCAGCTTCTCAATCTGTGAA 325

Db 61 ACCCAAGATACCTGACTCAGGATCTCAGCTTGGATTCAGCTTCTCAATCTGTGAA 120

QY 326 ATTACATCAAAACCCACCACTACTAGCTTCAACACACGAGATCAAGTCAACCTG 385

Db 121 ATTACATCAAAACCCACCACTACTAGCTTCAACACACGAGATCAAGTCAACCTG 180

QY 386 CAACCCACACAGTCAACATCAACCAACCAACCAACCAACCAACCAACCAACCAAC 445

Db 181 CAACCCACACAGTCAACATCAACCAACCAACCAACCAACCAACCAACCAACCAAC 240

QY 446 ACTCAAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505

Db 241 ACTCAAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300

QY 506 GTGTTAACTTTGTACCTGACATATGACAGCAACCAATCCAACTGCTGCTATCTGC 565

Db 301 GTGTTAACTTTGTACCTGACATATGACAGCAACCAATCCAACTGCTGCTATCTGC 360

QY 566 AAAAGATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625

Db 361 AAAAGATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420

QY 626 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 685

Db 421 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480

QY 686 ACCACCAAGCCACAGAAAGCCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 745

Db 481 ACCACCAAGCCACAGAAAGCCATCAACACCAACCAACCAACCAACCAACCAACCAACCA 540

QY 746 GTGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805

Db 541 GTGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600

QY 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCGAGACCCA 865

Db 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCGAGACCCA 660

QY 866 TCACAACCTCTCAATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920

Db 661 TCACAACCTCTCAATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715

RESULT 14

AAV38298 ID AAV38298 standard; DNA; 696 BP.

AAV38298; XX

26-OCT-1998 (first entry) XX

DT

```

XX Respiratory syncytial virus glycoprotein G gene portion.
DE
XX
KM RSV: glycoprotein G; pneumococcal surface protein A; PspA;
KM infection; Streptococcus pneumoniae; sepsis; otitis media;
KM meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;
KM ss.
XX Respiratory syncytial virus.
XX
XX W09824927-A1.
XX
XX 11-JUN-1998.
XX
XX 04-DEC-1997; 97MO-US22847.
XX
XX 04-DEC-1996; 96US-0759505.
XX
XX (UYAL-) UNIV ALABAMA.
XX
XX Briles DE, Curfiel DT, McDaniel LS;
XX
XX WPI; 1998-33343/29.
XX
XX Plasmid containing pneumococcal epitope for expression in eukaryotic
XX cells - useful for eliciting immunological response to pneumococcal
XX infection or sepsis
XX
XX Example 1; Fig 1C; 47pp; English.
XX
XX This is a portion of the respiratory syncytial virus glycoprotein G
XX (RSVg) gene. It has been inserted into plasmid pcDNA3 (see AAV38297)
XX to create plasmid pG74. This plasmid contains a human
XX cytomegalovirus immediate early promoter and the RSVg gene portion
XX such that when an in-frame fusion is made, the resultant fusion
XX protein may be transported to, and anchored in, a mammalian cell
XX membrane where it can be exposed to the host immune system.
XX Insertion of pneumococcal surface protein A (PspA) coding sequence
XX created plasmid pKSD2601. Intramuscular immunisation of BALB/c
XX mice with pKSD2601 induced protection against an otherwise lethal
XX challenge with a capsular type 3 pneumococcus. A claimed plasmid
XX for expression of pneumococcal epitope DNA in eukaryotic cells
XX includes a promoter for driving expression in a eukaryotic cell
XX (e.g. HCMV-IE), DNA encoding a leader sequence (e.g. of RSVg)
XX and DNA encoding a pneumococcal epitope such as PspA. The invention
XX also provides a vaccine comprising the plasmid and a suitable
XX carrier or diluent, and optionally one or more cytokines or DNA
XX encoding them, or a bacterial delivery system. The vaccine is used
XX to elicit an immunological response in a host, including humans,
XX susceptible to pneumococcal infection or sepsis. The plasmid can
XX also be used to express a pneumococcal epitope of interest in vitro.*
XX
XX Sequence 696 BP; 284 A; 211 C; 78 G; 123 T; 0 other;
XX
XX
XX Query Match 69.3%; Score 637.4; DB 19; Length 696;
XX Best Local Similarity 96.1%; Pred. No. 1.9e-143;
XX Matches 664; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
XX
XX
XX 8 ATGTCCAAAACAGACGACGACCGCTAAGACACTAGAAAAGCTGGAGCACTCTC 67
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 1 ATGTCCAAAACAGACGACGACCGCTAAGACACTAGAAAAGCTGGAGCACTCTC 60
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 68 AATCATTTATTTATTCATATCATCGGCTTATATTAAGTTAAATCTTAAATCTGTAGACAA 127
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 61 AATCATTTATTTATTCATATCATCGGCTTATATTAAGTTAAATCTTAAATCTGTAGACAA 120
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 128 ATGCATTATTCATTCGCAATGATTAATCTCACTTATATTAAGCAAGCATCTCTA 187
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 121 ATGCATTATTCATTCGCAATGATTAATCTCACTTATATTAAGCAAGCATCTCTA 180
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 188 TTCAATAGCCTGGGCAACCAAGTCACTAAGCACTGCAATCATCAAGATGCAACA 247
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX 181 TTCAATAGCCTGGGCAACCAAGTCACTAAGCACTGCAATCATCAAGATGCAACA 240
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

XX
XX 248 AGCCAGATTAACAAACAAACCCCAACATACCTCACTCAAGATCTGCACTTGGAAATCAGC 307
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 241 AGCCAGATTAACAAACAAACCCCAACATACCTCACTCAAGATCTGCACTTGGAAATCAGT 300
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 308 TTCTCCAAATCTGTGAAATTTACATCACAACCCACCACTACTGCTTCAAGTACACCA 367
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 301 CCTCTAATCCGCTGTGAATTTACATCACAACCCACCACTACTGCTTCAAGTACACCA 360
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 368 GGAGTCAAAGTCAAAACCTGCAACCCCAACAGTCAAGACTAAACACCAACAAAC- 426
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 361 GGAGTCAAAGTCAAAACCTGCAACCCCAACAGTCAAGACTAAACACCAACAAAC- 420
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 427 AACACAAACCCCAACAGCCCACTACAAAACAGCCCAACAAACCAACCAACAAAC- 486
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 421 AACACAAACCCCAACAGCCCACTACAAAACAGCCCAACAAACCAACCAACAAAC- 480
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 487 TAATGATTTTCACTTGAAGTGTTAACCTTGTACCTGCGAGCATATGACCAACCAATTC 546
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 481 TAATGATTTTCACTTGTGAAGTGTTAACCTTGTACCTGCGAGCATATGACCAACCAATTC 540
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 547 AACCTGTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCCAC 606
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 541 AACCTGTGGGCTATCTGCAAAAGATACCAACCAAAACCAAGGAAAGAAACCAACCCAC 600
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 607 CAAGGCTCAAAAAAAGCAACCTTCAAGACAAACCAAAACATCTCAACCTCTAAGCAC 666
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 601 CAAGGCTCAAAAAAAGCAACCTTCAAGACAAACCAAAACATCTCAACCTCTAAGCAC 660
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 667 TAAACCAAGCAAGTACCCACCAACAGCC 697
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX 661 TAAATCAAGCAAGTACCCACCAACAGCC 691
XX | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
XX
XX
XX RESULT 15
XX ID AAV17552 standard; cDNA: 15225 BP.
XX
XX AAV17552:
XX
XX 20-JUL-1998 (first entry)
XX
XX Respiratory syncytial virus genome.
XX
XX RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.
XX
XX Human respiratory syncytial virus B-1.
XX
XX W09802530-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97MO-US12269.
XX
XX 23-MAY-1997; 97US-0047634.
XX
XX 15-JUL-1996; 96US-0021773.
XX
XX 09-MAY-1997; 97US-0046141.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Bukreyev AA, Collins PL, Juhasz K, Murphy BR, Teng MN;
XX Whitehead SS;
XX
XX WPI; 1998-110579/10.
XX
XX Attenuated respiratory syncytial virus vaccines - useful to protect
XX individuals against RSV infection
XX
XX Example 4; Page 195-202; 238pp; English.
XX
XX This is the complete nucleotide sequence of the wild-type B-1
XX respiratory syncytial virus (RSV). The genome is negative-sense;
XX the 5'-3' positive-sense sequence of D46 is provided in AAV17553.
XX

```


GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:49:02 ; Search time 38.8257 Seconds
(without alignments)
7266.906 Million cell updates/sec

Title: US-09-462-816-1

Perfect score: 920
Sequence: 1 tgcacaactgtccaaaaa.....gtagtataaaaaaaaaa 920

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: Issued_Patents_NA.*
2: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/Backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	920	100.0	920	2	US-08-467-963C-7
2	920	100.0	920	2	US-08-838-189D-7
3	920	100.0	920	3	US-08-852-344D-7
4	920	100.0	920	3	US-08-344-639E-7
5	920	100.0	920	4	US-08-467-969A-7
6	920	100.0	920	4	US-08-467-961A-7
7	920	100.0	920	4	US-08-001-554A-7
8	894	97.2	894	2	US-08-467-963C-28
9	894	97.2	894	2	US-08-838-189D-28
10	894	97.2	894	2	US-08-852-344D-28
11	894	97.2	894	3	US-08-344-639E-28
12	853.4	92.8	15222	2	US-08-801-898A-23
13	853.4	92.8	15222	4	US-08-962-650-12
14	853.4	92.8	15223	2	US-08-892-403A-1
15	853.4	92.8	15223	4	US-08-720-132-1
16	441.4	48.0	15225	2	US-08-892-403A-2
17	212.8	28.1	1050	3	US-08-836-501-76
18	210.8	22.9	303	3	US-08-721-979A-1
19	210.8	22.9	303	3	US-08-836-504A-1
20	210.8	22.9	303	3	US-08-836-501-1
21	210.8	22.9	303	4	US-09-654-289-1
22	209.6	22.6	1071	3	US-08-836-501-77
23	207.6	22.6	303	3	US-08-721-979A-3
24	207.6	22.6	303	3	US-08-836-504A-3
25	207.6	22.6	303	4	US-08-836-501-3
26	207.6	22.6	303	4	US-09-654-289-3
27	201.2	21.9	303	3	US-08-721-979A-14

28	201.2	21.9	303	3	US-08-836-501-14	Sequence 14, Appl
29	201.2	21.9	303	4	US-09-654-289-14	Sequence 14, Appl
30	137	14.9	840	5	PCT-US91-08177-12	Sequence 12, Appl
31	137	14.9	7323	5	PCT-US91-08177-1	Sequence 1, Appl
32	128.6	14.0	183	3	US-08-721-979A-23	Sequence 23, Appl
33	128.6	14.0	183	3	US-08-836-501-23	Sequence 23, Appl
34	128.6	14.0	183	4	US-09-654-289-23	Sequence 23, Appl
35	126	13.7	303	3	US-08-721-979A-2	Sequence 2, Appl
36	126	13.7	303	3	US-08-836-501A-2	Sequence 2, Appl
37	126	13.7	303	4	US-09-654-289-2	Sequence 2, Appl
38	125.4	13.6	183	3	US-08-721-979A-30	Sequence 30, Appl
39	125.4	13.6	183	3	US-08-836-501-30	Sequence 30, Appl
40	125.4	13.6	183	4	US-09-654-289-30	Sequence 30, Appl
41	124.2	13.5	177	3	US-08-721-979A-24	Sequence 24, Appl
42	124.2	13.5	177	3	US-08-836-501-24	Sequence 24, Appl
43	124.2	13.5	177	4	US-09-654-289-24	Sequence 24, Appl
44	122.8	13.3	303	3	US-08-721-979A-4	Sequence 4, Appl
45						

ALIGNMENTS

RESULT 1
US-08-467-963C-7
Sequence 7, Application US/08467963C
Patent No. 5968776
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
APPLICANT: EMASYSYN, Mary E
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,963C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/838,189
FILING DATE: 16-Apr-1997
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-Jan-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-Jan-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-474 MTS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-963C-7

Query Match 100.0%; Score 920; DB 2; Length 920;
 Best Local Similarity 100.0%; Pred. No. 4,5e-243;
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 TGCACATGTCACAAAGAGACAGCCAGCGCTAGACACTGAGAAAGACCTGGGA 60
DB 1 TGCACATGTCACAAAGAGACAGCCAGCGCTAGACACTGAGAAAGACCTGGGA 60
QY 61 CACCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAATTTAAATCTGT 120
DB 61 CACCTCAATCATTTATTTATTCATATCATGCGGCTTATATTAATTTAAATCTGT 120
QY 121 AGCACAATGACATTTATTCATATTCATGCGAATGATATCTCACTTAAATTAACG 180
DB 121 AGCACAATGACATTTATTCATATTCATGCGAATGATATCTCACTTAAATTAACG 180
QY 181 CATCATTTATAGCTCGGCAAAACCAAGTACACTACAACTGCAATCTATACAGA 240
DB 181 CATCATTTATAGCTCGGCAAAACCAAGTACACTACAACTGCAATCTATACAGA 240
QY 241 TGCACAGCGCAGATCAAGAACACACCCACATACCTCTAGAGATCTGAGCTGG 300
DB 241 TGCACAGCGCAGATCAAGAACACACCCACATACCTCTAGAGATCTGAGCTGG 300
QY 301 AATCAGCTTCTCAATCTGTGAAATTTACATACAAACCCACACATACCTTCAAC 360
DB 301 AATCAGCTTCTCAATCTGTGAAATTTACATACAAACCCACACATACCTTCAAC 360
QY 361 AACACCGAGAGTCAAGTCAAACTGCAACCCCAACAGTAAACCTTAAACCAACAC 420
DB 361 AACACCGAGAGTCAAGTCAAACTGCAACCCCAACAGTAAACCTTAAACCAACAC 420
QY 421 AACCCAAACCAACCCAGACCCACTACAAACAAACGCGCAAAACCAACCAACAA 480
DB 421 AACCCAAACCAACCCAGACCCACTACAAACAAACGCGCAAAACCAACCAACAA 480
QY 481 ACCCAATATGATTTTCACTTCAAGTGTAACTTTGACCTGAGCATATGACGAA 540
DB 481 ACCCAATATGATTTTCACTTCAAGTGTAACTTTGACCTGAGCATATGACGAA 540
QY 541 CAATCCAACTGCTGGGCTATGCAAAAGATACAAACAAACCAACGAGAAAGAAAC 600
DB 541 CAATCCAACTGCTGGGCTATGCAAAAGATACAAACAAACCAACGAGAAAGAAAC 600
QY 601 CACCAACCAAGCTTACAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
DB 601 CACCAACCAAGCTTACAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
QY 661 AACCACTAAACCAAGAGATACCCACCAACCAAGCCCAAGAGCCCAACCAACAC 720
DB 661 AACCACTAAACCAAGAGATACCCACCAACCAAGCCCAAGAGCCCAACCAACAC 720
QY 721 CACCAAAACCAACATCACTACACTGCTACCAACCAACCAACCAACCAACCAAC 780
DB 721 CACCAAAACCAACATCACTACACTGCTACCAACCAACCAACCAACCAACCAAC 780
QY 781 CACAAGTCAATGGAACCTTCCACTACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840
DB 781 CACAAGTCAATGGAACCTTCCACTACCTCTCGAAGGCAATCTAAGCCCTTCTCA 840
QY 841 AGTCTCACAACATCGAGACACCATACAAACCTTCTGCAACCAACCAACAGCGCA 900
DB 841 AGTCTCACAACATCGAGACACCATACAAACCTTCTGCAACCAACCAACAGCGCA 900
QY 901 GTAGTTATTTAAAAAAA 920
DB 901 GTAGTTATTTAAAAAAA 920

```

RESULT 2
 US-08-838-189D-7
 ; Sequence 7, Application US/0838189D

```

; Patent No. 5998169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: ERASYSYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,189D
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-687 MJS:jb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 920 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-838-189D-7
;
; Query Match 100.0%; Score 920; DB 2; Length 920;
; Best Local Similarity 100.0%; Pred. No. 4,5e-243;
; Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

361 AACACGAGAGTCAAGTCAAACTGCAACCCACAACAGTCAAGACTAAATAACACACAGAC 420
|||||
361 AACACGAGAGTCAAGTCAAACTGCAACCCACAACAGTCAAGACTAAATAACACACAGAC 420
421 AACCCAAACACACCCAGCAGCCACTACAAAACAGCCAAACCAACCAACCAACCA 480
|||||
421 AACCCAAACACACCCAGCAGCCACTACAAAACAGCCAAACCAACCAACCAACCA 480
481 ACCCAATATATATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540
|||||
481 ACCCAATATATATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540
481 ACCCAATATATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 600
|||||
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 600
601 CACCAACAGCCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
|||||
601 CACCAACAGCCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
661 AACCACTAAACCAAGAGTACCCACCAAGCCCAAGAGGCAACCAACCAACCAAC 720
|||||
661 AACCACTAAACCAAGAGTACCCACCAAGCCCAAGAGGCAACCAACCAACCAAC 720
721 CACCAAAACCAACATCACAACCTGCTGACCAACCAACCAACCAACCAACCAAC 780
|||||
721 CACCAAAACCAACATCACAACCTGCTGACCAACCAACCAACCAACCAACCAAC 780
781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGCAATCAAGCCCTTCTCA 840
|||||
781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCGAAGCAATCAAGCCCTTCTCA 840
841 AGCTTCACACATTCGAGACACCAATCACAACCTCATCTGCAACCAACCAACCA 900
|||||
841 AGCTTCACACATTCGAGACACCAATCACAACCTCATCTGCAACCAACCAACCA 900
901 GTAGTATTATTAATAAAAAA 920
|||||
901 GTAGTATTATTAATAAAAAA 920
Db 901 GTAGTATTATTAATAAAAAA 920

RESULT 3
US-08-852-344D-7
Sequence 7, Application US/08852344D
Patent No. 6017539
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
APPLICANT: EWASYSHV, Mary E
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
TITLE OF INVENTION: SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,344D
FILING DATE: 07-MAY-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,639
FILING DATE: 14-NOV-1994

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-852-344D-7

Query Match 100.0%; Score 920; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 4,5e-243;
Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 TGCAAAATGTCACAAAACCAAGACCAAGCACCAGCTAAGACACTGAAAAGACCTGGGA 60
|||||
1 TGCAAAATGTCACAAAACCAAGACCAAGCACCAGCTAAGACACTGAAAAGACCTGGGA 60
61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATTAATGTTAAATCTATCTGT 120
|||||
61 CACTCTCAATCATTTATTTATTCATATCATGCGGCTTATTAATGTTAAATCTATCTGT 120
121 AGCACAATCATCATTTATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 180
|||||
121 AGCACAATCATCATTTATTCATATTCATATTCATATTCATATTCATATTCATATTCAT 180
181 CATCATATTCATTCAGCTTCGGCAACCAAGCAAGTACACTAAGCACTGCAATCTATACAGA 240
|||||
181 CATCATATTCATTCAGCTTCGGCAACCAAGCAAGTACACTAAGCACTGCAATCTATACAGA 240
241 TGCAACAAGCCAGATCAAGAACACACACCCACATACCTGATGATGATGATGATGATG 300
|||||
241 TGCAACAAGCCAGATCAAGAACACACACCCACATACCTGATGATGATGATGATGATG 300
301 AATGAGCTTCTCAATCTGCTGTAATTTACATCAACCAACCAACCAACCAACCAACCA 360
|||||
301 AATGAGCTTCTCAATCTGCTGTAATTTACATCAACCAACCAACCAACCAACCAACCA 360
361 AACACGAGAGTCAAGTCAAACTGCAACCCACAACAGTCAAGACTAAATAACACACAGAC 420
|||||
361 AACACGAGAGTCAAGTCAAACTGCAACCCACAACAGTCAAGACTAAATAACACACAGAC 420
421 AACCCAAACACACCCAGCAGCCACTACAAAACAGCCAAACCAACCAACCAACCAACCA 480
|||||
421 AACCCAAACACACCCAGCAGCCACTACAAAACAGCCAAACCAACCAACCAACCAACCA 480
481 ACCCAATATATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540
|||||
481 ACCCAATATATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540
481 ACCCAATATATTTTCACTTGAAGTGTAACTTTGTACCTGCGACATATGACGAA 540
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 600
|||||
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 600
541 CAATCCAACTGCTGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAAC 600
601 CACCAACAGCCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
|||||
601 CACCAACAGCCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
601 CACCAACAGCCTACAAAAAACCAACCTTCAAGACCAACCAACCAACCAACCTCA 660
661 AACCACTAAACCAAGAGTACCCACCAAGCCCAAGAGGCAACCAACCAACCAACCA 720
|||||
661 AACCACTAAACCAAGAGTACCCACCAAGCCCAAGAGGCAACCAACCAACCAACCA 720
721 CACCAAAACCAACATCACAACCTGCTGACCAACCAACCAACCAACCAACCAACCA 780
|||||
721 CACCAAAACCAACATCACAACCTGCTGACCAACCAACCAACCAACCAACCAACCA 780

QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
QY 841 AGTCTCCACACATCCGCGACCCATCAGCAACCTCTCATCTCCACCCACACAGCCCA 900
Db 841 AGTCTCCACACATCCGCGACCCATCAGCAACCTCTCATCTCCACCCACACAGCCCA 900
QY 901 GTAGTTATTAAAAA 920
Db 901 GTAGTTATTAAAAA 920

RESULT 4

US-08-344-639E-7
Sequence 7, Application US/08344639E
Patent No. 6033668
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasysbyn, Mary E
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
PROTECTION AGAINST PARAINFLUENZA VIRUS
TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th floor
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/344,639E
FILING DATE: 14-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-391 MIS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ. ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-344-639E-7

Query Match 100.0%; Score 920; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 4.5e-243;

Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 TGCACCAATGTCACCAACAGACCAAGCACCCTAGACACTAGAAAAGACCTGGGA 60
|||||

Db 1 TGCACCAATGTCACCAACAGACCAAGCACCCTAGACACTAGAAAAGACCTGGGA 60
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATTAAGTTAAATCTTAAATCTGT 120
Db 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTTATTAAGTTAAATCTTAAATCTGT 120
QY 121 AGCACAATACATTTATTCATTCCTGCGCAATGATATCTCAACTATTAATTAAGC 180
Db 121 AGCACAATACATTTATTCATTCCTGCGCAATGATATCTCAACTATTAATTAAGC 180
QY 181 CATCATATTCATATGCTCGGCAACCCAAAGTCACTATACACTGCAATGATCAAGA 240
Db 181 CATCATATTCATATGCTCGGCAACCCAAAGTCACTATACACTGCAATGATCAAGA 240
QY 241 TGCACCAAGCGATCAACCAACCAACCCCAATACCTCATCTGAGTATCTCCAGCTGG 300
Db 241 TGCACCAAGCGATCAACCAACCAACCCCAATACCTCATCTGAGTATCTCCAGCTGG 300
QY 301 AATCAGCTTCTCAATCTGTCTGAAATTAATCATCAGAAACCAACCATATAGCTTCAAC 360
Db 301 AATCAGCTTCTCAATCTGTCTGAAATTAATCATCAGAAACCAACCATATAGCTTCAAC 360
QY 361 AACACAGAGATCAATGCAAACTGCAACCCCAACAGTCAAGACTAAAAACACAAAC 420
Db 361 AACACAGAGATCAATGCAAACTGCAACCCCAACAGTCAAGACTAAAAACACAAAC 420
QY 421 AACCCAAACACCAACCCAGCAACCCACTACAAACCAACCCCAACCAACCAACCAAC 480
Db 421 AACCCAAACACCAACCCAGCAACCCACTACAAACCAACCCCAACCAACCAACCAAC 480
QY 481 ACCCAATTAATGATTTTCACTTGAAGTGTTAACCTTGTACCTGACCATATGACAGAA 540
Db 481 ACCCAATTAATGATTTTCACTTGAAGTGTTAACCTTGTACCTGACCATATGACAGAA 540
QY 541 CAATCAACCTGCTGGCTATCTGAAAGAAATACCAACCAACCAACCAACCAACCAAC 600
Db 541 CAATCAACCTGCTGGCTATCTGAAAGAAATACCAACCAACCAACCAACCAACCAAC 600
QY 601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCTCA 660
Db 601 CACCAACCAAGCTTACAAAAAACCAACCTTCAAGCAACCAACCAACCAACCAACCTCA 660
QY 661 AACCACTAAACCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCTCA 720
Db 661 AACCACTAAACCAAGGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCTCA 720
QY 721 CACCAAAACCAACATCAACATCACTGCTCACCACCAACCAACCAACCAACCAACCT 780
Db 721 CACCAAAACCAACATCAACATCACTGCTCACCACCAACCAACCAACCAACCAACCT 780
QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
Db 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAGGCAATCTAAGCCCTTCTCA 840
QY 841 AGTCTCCACACATCCGCGACCCATCAGCAACCTCTCATCTCCACCCACACAGCCCA 900
Db 841 AGTCTCCACACATCCGCGACCCATCAGCAACCTCTCATCTCCACCCACACAGCCCA 900
QY 901 GTAGTTATTAAAAA 920
Db 901 GTAGTTATTAAAAA 920

RESULT 5

US-08-467-969A-7
Sequence 7, Application US/08467969A
Patent No. 6168786
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasysbyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21

Page 5

OY	421	AACCCAAAGCACAACCAGCAACCCACATGACAAAACAAACGCCAATAAACAAACACACCAAAC	480
Dd	421	AACCCAAAGCACAACCAGCAACCCACATGACAAAACAAACGCCAATAAACAAACACACCAAAC	480
OY	481	ACCCAATTAATGATTTCATCTTGGAAGTGTTAACCTTTGTACCGTGAGCATATGACGAA	540
Dd	481	ACCCAATTAATGATTTCATCTTGGAAGTGTTAACCTTTGTACCGTGAGCATATGACGAA	540
OY	541	CAATCCAAACCTGTGGGCTATCTGCAAAAGAATACCAAAACAAAAAACAGAAAGAAAC	600
Dd	541	CAATCCAAACCTGTGGGCTATCTGCAAAAGAATACCAAAACAAAAAACAGAAAGAAAC	600
OY	601	CACACCAAGGCTACAAAAAAAACCAACTCTCAAGACACCAAAAAAAGATTCTCAAACTCA	660
Dd	601	CACACCAAGGCTACAAAAAAAACCAACTCTCAAGACACCAAAAAAAGATTCTCAAACTCA	660
OY	661	AACCACTAAACCAAGGAAGTAGCCACACCAAGCCACAGAAGCCACACATCAACAC	720
Dd	661	AACCACTAAACCAAGGAAGTAGCCACACCAAGCCACAGAAGCCACACATCAACAC	720
OY	721	CACCAAAACCAACATCACACTGCTCTCACCAACAACACACAGGAAGAAATCCAACACT	780
Dd	721	CACCAAAACCAACATCACACTGCTCTCACCAACAACACACAGGAAGAAATCCAACACT	780
OY	781	CACAAGTCAAATGGAAGAACCTTCACACTCAACTCTCTCCGAAGCAATCTAGCCCTTCTCA	840
Dd	781	CACAAGTCAAATGGAAGAACCTTCACACTCAACTCTCTCCGAAGCAATCTAGCCCTTCTCA	840
OY	841	AGTCTCCACACATCCGAGACCCATATACAGCCCTATCTCCACCCAAACATACAGCCA	900
Dd	841	AGTCTCCACACATCCGAGACCCATATACAGCCCTATCTCCACCCAAACATACAGCCA	900
OY	901	GTAAGTATTAAAAAAA 920	
Dd	901	GTAAGTATTAAAAAAA 920	

RESULT 6
US-08-467-961A-7
Sequence 7, Application US/08467961A
Patent No. 6171783
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewaszshyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6TH FLOOR
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,961A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-476 MIS:bn
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 TELEX: 065-24567 SIMBAS
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 920 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-467-961A-7

Query Match 100.0%; Score 920; DB 4; Length 920;
 Best Local Similarity 100.0%; Pred. No. 4.5e-243;
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACAAATGTCACAAAGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
 DB 1 TGCACAAATGTCACAAAGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
 QY 61 CACTCTCAATCAATTTATTTATTCATTCATCGGGCTTATTAAGTTAAATCTTGT 120
 DB 61 CACTCTCAATCAATTTATTTATTCATTCATCGGGCTTATTAAGTTAAATCTTGT 120
 QY 121 AGCACAATTCATATTCATTCATTCGCAATGATTAATTCATCTTACTTATTAACAGC 180
 DB 121 AGCACAATTCATATTCATTCATTCGCAATGATTAATTCATCTTACTTATTAACAGC 180
 QY 181 CATCATATTCATATTCATTCGCGCAACCAAGTCACTAATCACTGCAATCATACAGA 240
 DB 181 CATCATATTCATATTCATTCGCGCAACCAAGTCACTAATCACTGCAATCATACAGA 240
 QY 241 TGCACAAAGCCAGATCAAGACCAACCCCAATACCTCACTCAGATCTCAGCTTGG 300
 DB 241 TGCACAAAGCCAGATCAAGACCAACCCCAATACCTCACTCAGATCTCAGCTTGG 300
 QY 301 AATCAGCTCTCCAAATCTGCTGAATTTACATCAACCAACCACTACTGCTTCAAC 360
 DB 301 AATCAGCTCTCCAAATCTGCTGAATTTACATCAACCAACCACTACTGCTTCAAC 360
 QY 361 AACACGAGAGTCAAGTCAAACTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGC 420
 DB 361 AACACGAGAGTCAAGTCAAACTGCAACCAACAGTCAAGTCAAGTCAAGTCAAGC 420
 QY 421 AACCCAAACCAACCAAGAGCCCACTACAAAACCAAGCCCAAAACCAACCAACCA 480
 DB 421 AACCCAAACCAACCAAGAGCCCACTACAAAACCAAGCCCAAAACCAACCAACCA 480
 QY 481 ACCCAATATGATTTTCACTTCGAGTGTAACTTTGTAACCTTGAGCATATGACGAA 540
 DB 481 ACCCAATATGATTTTCACTTCGAGTGTAACTTTGTAACCTTGAGCATATGACGAA 540
 QY 541 CAATTCACCTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAGGAAAGAAAC 600
 DB 541 CAATTCACCTGCTGGGCTATCTGCAAAAGAAATACCAAAACCAAGGAAAGAAAC 600
 QY 601 CACACCAACGCTACAAAACCAACCTTCAAGACACCAAAAAGATCTCAAACTCA 660
 DB 601 CACACCAACGCTACAAAACCAACCTTCAAGACACCAAAAAGATCTCAAACTCA 660
 QY 661 AACCACTAAACCAAGAGATACCAACCAAGCCCAAGAGAGAGCAACCATCAACAC 720
 DB 661 AACCACTAAACCAAGAGATACCAACCAAGCCCAAGAGAGAGCAACCATCAACAC 720
 QY 721 CACCAAAACCAACATCAACATCACTGCTACCAACACCAAGCAAGGAAATCCAAACT 780
 DB 721 CACCAAAACCAACATCAACATCACTGCTACCAACACCAAGCAAGGAAATCCAAACT 780
 QY 781 CACAAGTCAATGGAACCTTCACCTCAACCTCTCGAAGGCAATCTAAGCCCTCTCA 840
 DB 781 CACAAGTCAATGGAACCTTCACCTCAACCTCTCGAAGGCAATCTAAGCCCTCTCA 840

DB 781 CACAAGTCAATGGAACCTTCACCTCAACCTCTCGAAGGCAATCTAAGCCCTCTCA 840
 QY 841 AGTCTCCACACATTCGAGACCAATCAGACACCTTATCTCCAAACCAACAGCGCA 900
 DB 841 AGTCTCCACACATTCGAGACCAATCAGACACCTTATCTCCAAACCAACAGCGCA 900
 QY 901 GTAGTTATTAATAAAAAAAAA 920
 DB 901 GTAGTTATTAATAAAAAAAAA 920

RESULT 7

US-08-001-554A-7
 Sequence 7, Application US/08001554A
 Patent No. 6225091

GENERAL INFORMATION:
 APPLICANT: Klein, Michel H
 APPLICANT: Du, Run-Pan
 APPLICANT: Ewasyschy, Mary E
 TITLE OF INVENTION: Chimeric Immunogens
 NUMBER OF SEQUENCES: 21
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Slim & McBurney
 STREET: 330 University Avenue, 6th Floor
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/001,554A
 FILING DATE: 06-JAN-1993
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-286
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:
 LENGTH: 920 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-001-554A-7

Query Match 100.0%; Score 920; DB 4; Length 920;
 Best Local Similarity 100.0%; Pred. No. 4.5e-243;
 Matches 920; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TGCACAAATGTCACAAAGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
 DB 1 TGCACAAATGTCACAAAGACCAAGCCGCTAAGACACTAGAAAAGACCTGGGA 60
 QY 61 CACTCTCAATCAATTTATTTATTCATTCATCGGGCTTATTAAGTTAAATCTTGT 120
 DB 61 CACTCTCAATCAATTTATTTATTCATTCATCGGGCTTATTAAGTTAAATCTTGT 120
 QY 121 AGCACAATTCATATTCATTCATTCGCAATGATTAATTCATCTTACTTATTAACAGC 180
 DB 121 AGCACAATTCATATTCATTCATTCGCAATGATTAATTCATCTTACTTATTAACAGC 180
 QY 181 CATCATATTCATATTCATTCGCGCAACCAAGTCACTAATCACTGCAATCATACAGA 240
 DB 181 CATCATATTCATATTCATTCGCGCAACCAAGTCACTAATCACTGCAATCATACAGA 240

Db 601 AAGCTCAAAAAAACCACTTCAAGACAAACAAAAAGATCTCAACCTCAACACT 660
QY 668 AACCAAGAGATACCCACCACCAAGCCACAGAGAGCCAAACATCAACACCACAA 727
Db 661 AACCAAGAGATACCCACCACCAAGCCACAGAGAGCCAAACATCAACACCACAA 720
QY 728 ACAACATCACTACCTGCTCAGCAACACAGAGCAAGTCAAAATCTCAAGT 787
Db 721 ACAACATCACTACCTGCTCAGCAACACAGAGCAAGTCAAAATCTCAAGT 780
QY 788 CAATGGAACCTTCACTCACTCACTCCGAGAGCAATTAAGCTTCTCAAGTCTC 847
Db 781 CAATGGAACCTTCACTCACTCACTCCGAGAGCAATTAAGCTTCTCAAGTCTC 840
QY 848 ACAACATCCGAGACCCATCAACACCTCTATCTCCACCCCAACAGACGCCAG 901
Db 841 ACAACATCCGAGACCCATCAACACCTCTATCTCCACCCCAACAGACGCCAG 894

RESULT 9

US-08-838-189D-28
; Sequence 28, Application US/08838189D
; Patent No. 5998169
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EMASYSYN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/838,189D
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jfb
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-838-189D-28

Query Match 97.2%; Score 894; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 6e-236;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATGTCAAAAAACAGGACCAAGCCGCTAAGACACTAGAAAAGACCTGGACACTCTC 67
Db 1 ATGTCAAAAAACAGGACCAAGCCGCTAAGACACTAGAAAAGACCTGGACACTCTC 60
QY 68 AATCATTTATTTATTCATATCATCGGGCTTATATAGTTAAATCTTAAATCTGTAGCA 127
Db 61 AATCATTTATTTATTCATATTCATCGGGCTTATATAGTTAAATCTTAAATCTGTAGCA 120
QY 128 ATCAGATTATTCATTCCTGGCAATGATATATCTCACTTACTTATTAATATAGCCATATA 187
Db 121 ATCAGATTATTCATTCCTGGCAATGATATATCTCACTTACTTATTAATATAGCCATATA 180
QY 188 TTCATAGCTCGGCAAAACACAAAGTCACTAACTAACTGCAATCATACAGATGACACA 247
Db 181 TTCATAGCTCGGCAAAACACAAAGTCACTAACTAACTGCAATCATACAGATGACACA 240
QY 248 AGCCAGATCAGAGACACACACCCCAACATACCTCACTAGAGTCTCAGCTTGGATAGC 307
Db 241 AGCCAGATCAGAGACACACACCCCAACATACCTCACTAGAGTCTCAGCTTGGATAGC 300
QY 308 TTCTCAATCTGTCTGAATTTACATCAAAACGACACCTAGCTTCAACACACGA 367
Db 301 TTCTCAATCTGTCTGAATTTACATCAAAACGACACCTAGCTTCAACACACGA 360
QY 368 GGAGTCAAGTCAAAACCTGCAACCCACACAGTCAAGTCAAAACACACACACCA 427
Db 361 GGAGTCAAGTCAAAACCTGCAACCCACACAGTCAAGTCAAAACACACACCA 420
QY 428 ACACAAACCCAGAACCCCACTACAAACAAAGCCCAAAACCAACCAACCAACCAAT 487
Db 421 ACACAAACCCAGAACCCCACTACAAACAAAGCCCAAAACCAACCAACCAAT 480
QY 488 AATGATTTTCACTTGAAGTGTAACTTTTACCTTGCACATATGAGCAGCAATCA 547
Db 481 AATGATTTTCACTTGAAGTGTAACTTTTACCTTGCACATATGAGCAGCAATCA 540
QY 548 ACTCTGTGGCTATGTGCAAAAAGATACCAACAAACCAAGAGAGAGAGAGAGAGAG 607
Db 541 ACTCTGTGGCTATGTGCAAAAAGATACCAACAAACCAAGAGAGAGAGAGAGAGAG 600
QY 608 AAGCTTCAAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAACCTCAACCACT 667
Db 601 AAGCTTCAAAAAACCAACCTTCAAGACAAACCAAAAGATCTCAACCTCAACCACT 660
QY 668 AACCAAGAGATACCCACCACCAAGCCACAGAGAGCCACACATCAACACCAACAA 727
Db 661 AACCAAGAGATACCCACCACCAAGCCACAGAGAGCCACACATCAACACCAACAA 720
QY 728 ACAACATCACTACCTGCTCAGCAACACAGAGCAAGTCAAAATCTCAAGT 787
Db 721 ACAACATCACTACCTGCTCAGCAACACAGAGCAAGTCAAAATCTCAAGT 780
QY 788 CAATGGAACCTTCACTCACTCACTCCGAGAGCAATTAAGCTTCTCAAGTCTC 847
Db 781 CAATGGAACCTTCACTCACTCACTCCGAGAGCAATTAAGCTTCTCAAGTCTC 840
QY 848 ACAACATCCGAGACCCATCAACACCTCTATCTCCACCCCAACAGACGCCAG 901
Db 841 ACAACATCCGAGACCCATCAACACCTCTATCTCCACCCCAACAGACGCCAG 894

RESULT 10

US-08-852-344D-28
; Sequence 28, Application US/08852344D
; Patent No. 6017539
; GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; APPLICANT: EMASYSYN, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
; TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
; TITLE OF INVENTION: SYNCYTIAL VIRUS
; NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 6th Floor, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/852,344D
 FILING DATE: 07-MAY-1997
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/344,639
 FILING DATE: 14-NOV-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9200117.1
 FILING DATE: 06-JAN-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: STEWART, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-688 MIS:jfb
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 894 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-852-344D-28

Query Match 97.2%; Score 894; DB 3; Length 894;
 Best Local Similarity 100.0%; Pred. No. 6e-236;
 Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 ATGTCCAAAACAGAGCAACGCGGCTTAGACACTGAAAGAGCGGACACTCTC 67
 DB 1 ATGTCCAAAACAGAGCAACGCGGCTTAGACACTGAAAGAGCGGACACTCTC 60
 QY 68 AATCATTTTATTTATTCATATCATCGGCTTATTAAGTTAAATCTTAACTGACACAA 127
 DB 61 AATCATTTTATTTATTCATATCATCGGCTTATTAAGTTAAATCTTAACTGACACAA 120
 QY 128 ATCATTTTATTCATATTCATCGCAATGATATCTGAATCTTAAATTAATACAGCATCAT 187
 DB 121 ATCATTTTATTCATATTCATCGCAATGATATCTGAATCTTAAATTAATACAGCATCAT 180
 QY 188 TTTCATAGCTGGCAACGCAAGGACACACACACACACACACACACACACACACAC 247
 DB 181 TTTCATAGCTGGCAACGCAAGGACACACACACACACACACACACACACACACAC 240
 QY 248 ACCGAGATCAAGAC 307
 DB 241 ACCGAGATCAAGAC 300
 QY 308 TTCTCCAACTCTGTGAATTTATCATCAAAACACACACACACACACACACACAC 367
 DB 301 TTCTCCAACTCTGTGAATTTATCATCAAAACACACACACACACACACACACAC 360
 QY 368 GGAGTCAGTCAAACTGCAACGCGGACAGTCAGATCAAAACACACACACACAC 427
 DB 361 GGAGTCAGTCAAACTGCAACGCGGACAGTCAGATCAAAACACACACACACAC 420
 QY 428 ACACAACCCAGCAGCCCACTACAAACACAGCCCAACACACACACACACACAC 487
 DB 421 ACACAACCCAGCAGCCCACTACAAACACAGCCCAACACACACACACACAC 480

QY 488 AATGATTTTCACTTGAAGTGTTTTACTTTGATCCCTGACAGATATGAG::AACJATCCA 547
 DB 481 AATGATTTTCACTTGAAGTGTTTTACTTTGATCCCTGACAGATATGAG::AATCCA 540
 QY 548 ACCTGCTGGGCTATCTGCAAAAAGAAATACCAAAAACCAAGGAAAGAA::AGCAGC 607
 DB 541 ACCTGCTGGGCTATCTGCAAAAAGAAATACCAAAAACCAAGGAAAGAA::AGCAGC 600
 QY 608 AAGCTTACAAAAAACAACCTCTCAAGACACACAAAAAGATCTCAAACTCAAACTCA 667
 DB 601 AAGCTTACAAAAAACAACCTCTCAAGACACACAAAAAGATCTCAAACTCAAACTCA 660
 QY 668 AAACCAAGAAAGTACCCACACCAAGCCACAGAGAGGACCAACATCAACACACACAA 727
 DB 661 AAACCAAGAAAGTACCCACACCAAGCCACAGAGAGGACCAACATCAACACACAA 720
 QY 728 ACAACATCACAACCTGCTGACACACACACACACACACACACACACACACAC 787
 DB 721 ACAACATCACAACCTGCTGACACACACACACACACACACACACACACACAC 780
 QY 788 CAAATGGAACCTTCCACTCAACCTCTCGAGAGGCAATCAAGCCCTTCAAGTCTCC 847
 DB 781 CAAATGGAACCTTCCACTCAACCTCTCGAGAGGCAATCAAGCCCTTCAAGTCTCC 840
 QY 848 ACAACATCAGACACCCATCACAACCCATCTCCACCCACACACACACACAC 901
 DB 841 ACAACATCAGACACCCATCACAACCCATCTCCACCCACACACACACACAC 894

RESULT 11
 US-08-344-639E-28
 Sequence 28, Application US/08344639E
 Patent No. 6033668
 GENERAL INFORMATION:
 APPLICANT: Klein, Michel H
 APPLICANT: Du, Run-Pan
 TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
 TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
 TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
 NUMBER OF SEQUENCES: 38
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 330 University Avenue, 6th Floor
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/344,639E
 FILING DATE: 14-NOV-1994
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/001,554
 FILING DATE: 06-JAN-1993
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 9200117.1
 FILING DATE: 06-JAN-1992
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-391 MIS
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 TELLEX: 065-24567 STMBAS

INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-344-639E-28

Query Match 97.2%; Score 894; DB 3; Length 894;
Best Local Similarity 100.0%; Pred. No. 6e-236;
Matches 894; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 8 ATGTCCAAAACAGAGCAACGACCGGTAAAGACTGAGAAAGACCTGGGACACTCTC 67
Db 1 ATGTCCAAAACAGAGCAACGACCGGTAAAGACTGAGAAAGACCTGGGACACTCTC 60
QY 68 AATCATTTTATTTATTCATATCATCGGGCTATATTAAGTTAAATCTTAATCTGTAGACAA 127
Db 61 AATCATTTTATTTATTCATATCATCGGGCTATATTAAGTTAAATCTTAATCTGTAGACAA 120
QY 128 ATCACTATTCATCTTGTGGCAATGATATCTCACTTCACTTATATTTACGCCATCATA 187
Db 121 ATCACTATTTATTCATCTTGTGGCAATGATATCTCACTTCACTTATATTTACGCCATCATA 180
QY 188 TTCTATAGCTCGGCAACCAAGTCACTAACAAGTCAATGATCAATGATGCAACA 247
Db 181 TTCTATAGCTCGGCAACCAAGTCACTAACAAGTCAATGATCAATGATGCAACA 240
QY 248 AGCCAGATCAAGAACACACACCCCAATACCTCTCAGAGATCTCAGCTTGGATCAGC 307
Db 241 AGCCAGATCAAGAACACACCCCAATACCTCTCAGAGATCTCAGCTTGGATCAGC 300
QY 308 TTCTCCAACTGTCTGTAATTTATCATCAACAACCCACCATCTAGCTTCAACACACCA 367
Db 301 TTCTCCAACTGTCTGTAATTTATCATCAACAACCCACCATCTAGCTTCAACACACCA 360
QY 368 GGAGTCAAGTCAACCTGCAACCCCAAGTCAAGACTTAAACCAACCAACCCCAA 427
Db 361 GGAGTCAAGTCAACCTGCAACCCCAAGTCAAGACTTAAACCAACCAACCCCAA 420
QY 428 ACACAAACCCGCAAGGCCACTATACAAAACAGCCAAACCAACCAACCAACCCAA 487
Db 421 ACACAAACCCGCAAGGCCACTATACAAAACAGCCAAACCAACCAACCAACCCAA 480
QY 488 AATGATTTTCACTTCGAAAGTGTAACTTTGTACCTGACGATATGACGCAACAATCA 547
Db 481 AATGATTTTCACTTCGAAAGTGTAACTTTGTACCTGACGATATGACGCAACAATCA 540
QY 548 ACCTGCTGGCTATCTGCAAAAGATTCACAACAAAACAGAGAAAGAAACCCACACC 607
Db 541 ACCTGCTGGCTATCTGCAAAAGATTCACAACAAAACAGAGAAAGAAAGAAACCCACACC 600
QY 608 AAGGCTTCAAAAACCAACCTTCAAGACAAACAAAAGATCTCAAACTCTCAACCACT 667
Db 601 AAGGCTTCAAAAACCAACCTTCAAGACAAACAAAAGATCTCAAACTCTCAACCACT 660
QY 668 AAAACCAAGAGATACCCACACCAAGCCCAAGAGAGGCAACCATCAACACACCAAA 727
Db 661 AAAACCAAGAGATACCCACACCAAGCCCAAGAGAGGCAACCATCAACACACCAAA 720
QY 728 ACAAACTCACAAGTACGCTGCTCAGCAACAACCAAGAAATCCAAAACCTCAACAAG 787
Db 721 ACAAACTCACAAGTACGCTGCTCAGCAACAACCAAGAAATCCAAAACCTCAACAAG 780
QY 788 CAATGGAACCTTCCACTACCTGCTCGAAGGCAATTAAGCCCTTCTCAAGTCTGC 847
Db 781 CAATGGAACCTTCCACTACCTGCTCGAAGGCAATTAAGCCCTTCTCAAGTCTGC 840
QY 848 ACAACATCGAGCAACCCATCAACAACCTCATCTCCACCAACCAACAGCCGAG 901
Db 841 ACAACATCGAGCAACCCATCAACAACCTCATCTCCACCAACCAACAGCCGAG 894
```

RESULT 12
US-08-801-898A-23
Sequence 23, Application US/08801898A
Patent No. 5998602

GENERAL INFORMATION:
APPLICANT: Torrence, Paul F.
APPLICANT: Silverman, Robert H.
APPLICANT: Cirino, Nick M.
APPLICANT: Li, Gulying
APPLICANT: Xiao, Wei
TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE
TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM type: Diskette

COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/801,898A
FILING DATE: 18-FEB-1997

CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:

NAME: Poissant, Brian M.
REGISTRATION NUMBER: 28,462

REFERENCE/DOCKET NUMBER: 8656-009
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741

TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:
LENGTH: 15222 base pairs

TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: DNA

FEATURE:
NAME/KEY: RSV-A2

LOCATION: 1...15222
OTHER INFORMATION:

US-08-801-898A-23

Query Match 92.8%; Score 853.4; DB 2; Length 15222;
Best Local Similarity 95.5%; Pred. No. 2.9e-224;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
QY 1 TCCAAACATGTCGCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 60
Db 4681 TCCAAACATGTCGCAAAACAGGACCAACGACCGCTAAGACACTAGAAAAGACCTGGGA 4740
QY 61 CACTCTCAATCTTATTTATTCATATCATCGGGCTTATTAAGTTAAATCTTAATCTGT 120
Db 4741 CACTCTCAATCTTATTTATTCATATCATCGGGCTTATTAAGTTAAATCTTAATCTGT 4800
QY 121 AGCAACAATACATTCATCTTGTGGCAATGATATCTCAAACTTCACTTATTAATTCAGC 180
Db 4801 AGCAACAATACATTCATCTTGTGGCAATGATATCTCAAACTTCACTTATTAATTCAGC 4860
QY 181 CATCATATTCATAGGCTCGGCAACCAAGTCACTAACAAGTCAATCAATGCAATACAGA 240
Db 4861 CATCATATTCATAGGCTCGGCAACCAAGTCACTAACAAGTCAATCAATGCAATACAGA 4920
QY 241 TGCACAGCCGATCAAGAACCAACCCCAACATACCTCACTGAGATCTCAGCTTGG 300
Db 4921 TGCACAGCCGATCAAGAACCAACCCCAACATACCTCACTGAGATCTCAGCTTGG 4980
```

QY 301 AATGAGCTTCTCCAAATCTGTCTGAAATTTACATCACAACCAACCCACTACTAGCTTCAAC 360
 Db 4981 AATGAGCTTCTCTATATCGTCTGAAATTTACATCACAACCAACCCACTACTAGCTTCAAC 5040
 QY 361 AAGACGAGAGTCAAGTCAAACTGCAACCCACAGCAAGTCAAGCTTAAAAACAACCAAC 420
 Db 5041 AACACGAGAGTCAAGTCAAACTGCAACCCACAGCAAGTCAAGCTTAAAAACAACCAAC 5100
 QY 421 AACCAACGACACCCAGCAAGCCCTACACAAAACAGCCAAAACCAACCAACCAACCA 480
 Db 5101 AACCTCAACGACACCCAGCAAGCCCTACACAAAACAGCCAAAACCAACCAACCAACCA 5160
 QY 481 ACCCAATATGATTTTCACTTCAAGTGTGTTAACTTTGCTGCTGAGCTATGACGCA 540
 Db 5161 ACCCAATATGATTTTCACTTCAAGTGTGTTAACTTTGCTGCTGAGCTATGACGCA 5220
 QY 541 CAATCCAAACCTGCTGGGCTATCTGCAAAAGATTCGAACCAATTCGAACCAACCA 600
 Db 5221 CAATCCAAACCTGCTGGGCTATCTGCAAAAGATTCGAACCAATTCGAACCAACCA 5280
 QY 601 CACCAACGAGCTTCAAAAAACCAACCTTCAAGACACCAACCAACCAACCAACCAAC 660
 Db 5281 CACCAACGAGCTTCAAAAAACCAACCTTCAAGACACCAACCAACCAACCAACCAAC 5340
 QY 661 AACCACTAAACGAAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCA 720
 Db 5341 AACCACTAAACGAAAGAGTACCCACCAACCAACCAACCAACCAACCAACCAACCA 5400
 QY 721 CACCAAAACCAACATCAGCTACACTGCTGACCAACCAACCAACCAACCAACCAAC 780
 Db 5401 CACCAAAACCAACATCAGCTACACTGCTGACCAACCAACCAACCAACCAACCAAC 5460
 QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 840
 Db 5461 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCA 5520
 QY 841 AGTCTCACAACATCCGAGCAGCCATACAAACCTTATCTCAGCCCAACCAACCAAC 900
 Db 5521 AGTCTCACAACATCCGAGCAGCCATACAAACCTTATCTCAGCCCAACCAACCAAC 5580
 QY 901 GTAGTTATTTAAAAAAA 919
 Db 5581 GTAGTTATTTAAAAACATA 5599

RESULT 13
 US-08-962-690-12/C
 ; Sequence 12, Application US/08962690
 ; Patent No. 6214805
 ; GENERAL INFORMATION:
 ; APPLICANT: Torrence, Paul F.
 ; APPLICANT: Silverman, Robert H.
 ; APPLICANT: Clirino, Nick M.
 ; APPLICANT: Li, Guiying
 ; APPLICANT: Xiao, Wei
 ; APPLICANT: Player, Mark R.
 ; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE OLIGONUCLEOTIDES
 ; TITLE OF INVENTION: EFFECTIVE TO TREAT RSV INFECTIONS
 ; FILE REFERENCE: 8656-019
 ; CURRENT APPLICATION NUMBER: US/08/962,690
 ; EARLIER APPLICATION NUMBER: 08/801,896
 ; EARLIER FILING DATE: 1997-02-14
 ; EARLIER APPLICATION NUMBER: 60/011,725
 ; EARLIER FILING DATE: 1996-02-15
 ; NUMBER OF SEQ ID NOS: 40
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 12
 ; LENGTH: 15222
 ; TYPE: DNA
 ; ORGANISM: respiratory syncytial virus
 US-08-962-690-12

Query Match 92.88; Score 853.4; DB 4; Length 15222;
 Best Local Similarity 95.58; Pred. No. 2,9e-224;
 Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;
 QY 1 TGCACATGTCGCAAAAAAGAGCCAGCCAGCCGTAAGACACTGAAAAAGCCGCGGA 60
 Db 10542 TGCACATGTCGCAAAAAAGAGCCAGCCAGCCGTAAGACACTGAAAAAGCCGCGGA 10483
 QY 61 CACTCTCAATCTTATTTATTCATATCATGCGCTTATATTAATTAATCTTAACTCTGT 120
 Db 10482 CACTCTCAATCTTATTTATTCATATCATGCGCTTATATTAATTAATCTTAACTCTGT 10423
 QY 121 AGCAAAATCATCATTTATTCATTTGCGCAATGATTAATTCGCAACTTCACTTA 180
 Db 10422 AGCAAAATCATCATTTATTCATTTGCGCAATGATTAATTCGCAACTTCACTTA 10363
 QY 181 CATCATATTCATGCTGCGCAACCAAGTCACAGTACAGTACAGTACAGTACAGTAC 240
 Db 10362 CATCATATTCATGCTGCGCAACCAAGTCACAGTACAGTACAGTACAGTACAGTAC 10303
 QY 241 TGCACAAAGCCAGATCAAGAACCAACCCCACTACTCTCAGATCTCAGCTTGG 300
 Db 10302 TGCACAAAGCCAGATCAAGAACCAACCCCACTACTCTCAGATCTCAGCTTGG 10243
 QY 301 AATGAGCTTCTCCATCTGTCTGAAATTTACATCACAACCAACCAACCAACCA 360
 Db 10242 AATGAGCTTCTCCATCTGTCTGAAATTTACATCACAACCAACCAACCAACCA 10183
 QY 361 AACACGAGAGTCAAGTCAACCTGCAACCCACAGTCACAGTCAAGTCAAGTCAAG 420
 Db 10182 AACACGAGAGTCAAGTCAACCTGCAACCCACAGTCACAGTCAAGTCAAGTCAAG 10123
 QY 421 AACCAAAACCAACCCAGCAAGCCCACTCAAAAACCAACCCCAACCAACCAACCA 480
 Db 10122 AACCAAAACCAACCCAGCAAGCCCACTCAAAAACCAACCCCAACCAACCAAC 10063
 QY 481 ACCCAATATGATTTTCACTTCAAGTGTGTTAACTTTGACCTGAGCATATGACGA 540
 Db 10062 ACCCAATATGATTTTCACTTCAAGTGTGTTAACTTTGACCTGAGCATATGACGA 10003
 QY 541 CAATCCAAACCTGCTGGGCTATCTGCAAAAGATTCGAACCAACCAACCAACCA 600
 Db 10002 CAATCCAAACCTGCTGGGCTATCTGCAAAAGATTCGAACCAACCAACCAACCA 5943
 QY 601 CACGACGAGCTTACAAAAACCAACCTTCAAGACACCAAAAAAGATTCGAAACCTCA 660
 Db 9942 CACTACCAACCCCAAAAAACCAACCTTCAAGACACCAAAAAAGATTCGAAACCTCA 9883
 QY 661 AACCACTAAACCAAGAGTACCCACGACCAAGCCCAAGAGGCGCAACCATCAACAC 720
 Db 9882 AACCACTAAACCAAGAGTACCCACGACCAAGCCCAAGAGGCGCAACCATCAACAC 9823
 QY 721 CACCAAAACCAACATCATCACTACACTGCTCAACCAACCAACCAACCAACCA 780
 Db 9822 CACCAAAACCAACATCATCACTACACTGCTCAACCAACCAACCAACCAACCAAC 9763
 QY 781 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTA 840
 Db 9762 CACAAGTCAATGGAACCTTCCACTCAACCTCTCCGAAGGCAATCTA 9703
 QY 841 AGTCTCACAACATCCGAGCAGCCATACAAACCTTATCTCAGCCCAACCAACCAAC 900
 Db 9702 AGTCTCACAACATCCGAGCAGCCATACAAACCTTATCTCAGCCCAACCAACCAAC 9643
 QY 901 GTAGTTATTTAAAAAAA 919
 Db 9642 GTAGTTATTTAAAAACATA 9624

RESULT 14
 US-08-892-403A-1
 ; Sequence 1, Application US/08892403A

APPLICATION NUMBER: US 60/007,083
FILING DATE: 27-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Parmelee, Steven W.
REGISTRATION NUMBER: 31,990
REFERENCE/DOCKET NUMBER: 15280-250-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-467-9600
TELEFAX: 415-576-0300
INFORMATION FOR SEQ. ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 15223 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other
US-08-720-132-1

Query Match 92.8%; Score 853.4; DB 4; Length 15223;
Best Local Similarity 95.5%; Pred. No. 2.9e-224;
Matches 878; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 1 TCGAACATGTGCAAAAACAGAGACGACGCGCTAAGACCTAGAAAAAGACCTGGGA 60
DB 4682 TCGAACATGTGCAAAAACAGAGACGACGCGCTAAGACCTAGAAAAAGACCTGGGA 4741
QY 61 CACTCTCAATCATTTATTTATTCATATCATCGGGCTATATAGTTAATCTTAATCTGT 120
DB 4742 CACTCTCAATCATTTATTTATTCATATCATCGGGCTATATAGTTAATCTTAATCTGT 4801
QY 121 AGCACAATTCATTTATTCATTTCTGGCAATGATTAATCTCACTTCACTTATATTACAGC 180
DB 4802 AGCACAATTCATTTATTCATTTCTGGCAATGATTAATCTCACTTCACTTATATTACAGC 4861
QY 181 CATCATATTCATAGCGCTGGCAAAACCAAACTGACACTAAGACGCAATCATACAGA 240
DB 4862 CATCATATTCATAGCGCTGGCAAAACCAAACTGACACTAAGACGCAATCATACAGA 4921
QY 241 TCGAACAGCCAGATCAGACAGACCAACCCCAATACCTCACTAGATCTCAGCTTGG 300
DB 4922 TCGAACAGCCAGATCAGACAGACCAACCCCAATACCTCACTAGATCTCAGCTTGG 4981
QY 301 AATCAGCTTCTCAATCTCTCTGAAATTCATACACAACCCACACCATTAAGCTTCAAC 360
DB 4982 AATCAGCTTCTCAATCTCTCTGAAATTCATACACAACCCACACCATTAAGCTTCAAC 5041
QY 361 AACACGAGAGTCAAGTCAAACTGCAACCCCAAGTCAAGTCAAAACCAACCAAC 420
DB 5042 AACACGAGAGTCAAGTCAAACTGCAACCCCAAGTCAAGTCAAAACCAACCAAC 5101
QY 421 AACCCAAACACACCCAGCAGCAGCTACAAACAGCCCAAAACCAAAACCAACCA 480
DB 5102 AACCTCAACACCAACCCAGCAGCAGCTACAAACAGCCCAAAACCAAAACCAACCA 5161
QY 481 ACCCAATATGATTTCTCACTTGAAGTGTTAACCTTGACCTCGACATATGCAGCA 540
DB 5162 ACCCAATATGATTTCTCACTTGAAGTGTTAACCTTGACCTCGACATATGCAGCA 5221
QY 541 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGAAAGAAC 600
DB 5222 CAATCCAACTGCTGGGCTATCTGCAAAAGATACCAAAACCAAAACCAAGAAAGAAC 5281
QY 601 CACCAACCAAGCTACAAAAAACCAACTTCAGAGCAACCAAAAAAGATCTCAAACTCA 660
DB 5282 CACTACCAAGCCCAAAAAAACCAACCTCAAGCAACCAAAAAAGATCTCAAACTCA 5341
QY 661 AACCACTAAACCAAGAGTACCCACACCAAGCCCAAGAGCCCAACCATCAACAC 720
DB 5342 AACCACTAAATCAAGAGATACCCACACCAAGCCCAAGAGCCCAACCATCAACAC 5401
QY 721 CACCAAAACCAACATCACACTACTGCTGACCAACCAACCAACAGAGAAATCCAAAAC 780
DB 5402 CACCAAAACCAACATCACACTACTGCTGACCAACCAACCAACAGAGAAATCCAAAAC 5461

QY 781 CACAAGTCAATGAAACCTTCCACTCAACCTCTCCGAAAGCAATCTTAAGCCTTTCTCA 840
DB 5462 CACAAGTCAATGAAACCTTCCACTCAACCTCTCCGAAAGCAATCTTAAGCCTTTCTCA 5521
QY 841 AGCTTCCACACATCCGAGCACCACATCAACCTCTCATCTGACCCAAACACAGCGCA 900
DB 5522 AGCTTCCACACATCCGAGCACCACATCAACCTCTCATCTGACCCAAACACAGCGCA 5581
QY 901 GTAGTTATTAATAAAAAA 919
DB 5582 GTAGTTACTTAATAAACATA 5600

Search completed: May 1, 2003, 12:35:08
Job time : 63.8257 secs


```

? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annonach Sequence Listing Engine vers. 1.1
? SEQ ID NO 20241
? LENGTH: 1635
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AC006547.9
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
? OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
? OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 8.6
? OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 9
? OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 12
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864 -761-20241

```

Query Match	8.5%;	Score 78.2;	DB 10;	Length 1635;
Best Local Similarity	46.6%;	Pred. No. 2.3e-10;		
Matches 354;	Conservative 0;	Mismatches 398;	Indels 7;	Gaps 3;

[illegible]

QY 847 CACAACATCCGAGCACCATTACAAACCTTATCTTCACC 865
+ + + + +
Db 1091 ACAGCACCAACCAACCATCTACTACCAACCAACCAACC 1129

RESULT 5
US-09-864-761-3471
Sequence 3471, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIORITY APPLICATION NUMBER: US 60/180,312
PRIORITY FILING DATE: 2000-02-04
PRIORITY APPLICATION NUMBER: US 60/207,456
PRIORITY FILING DATE: 2000-05-26
PRIORITY APPLICATION NUMBER: US 09/632,366
PRIORITY FILING DATE: 2000-08-03
PRIORITY APPLICATION NUMBER: GB 24263.6
PRIORITY FILING DATE: 2000-10-04
PRIORITY APPLICATION NUMBER: US 60/236,359
PRIORITY FILING DATE: 2000-09-27
PRIORITY APPLICATION NUMBER: PCT/US01/00666
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00667
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00664
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00669
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00665
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00668
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00663
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00662
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00661
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: PCT/US01/00670
PRIORITY FILING DATE: 2001-01-30
PRIORITY APPLICATION NUMBER: US 60/234,687
PRIORITY FILING DATE: 2000-09-21
PRIORITY APPLICATION NUMBER: US 09/608,408
PRIORITY FILING DATE: 2000-06-30
PRIORITY APPLICATION NUMBER: US 09/774,203
PRIORITY FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 3471
LENGTH: 1973
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006547.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12

US-09-864-761-3471

Query Match	8.5%;	Score 78.2;	DB 10;	Length 1973;
Best Local Similarity	46.6%;	Pred. No. 2.5e-10;		
Matches 354;	Conservative 0;	Mismatches 398;	Indels 7;	Gaps 3;

[illegible]

```

RESULT 6
US-09-878-574-4296/C
: Sequence 4296, Application US/09878574
: Patent No. US20020110548A1
:
GENERAL INFORMATION:
:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/333,535

```

```

: PRIOR FILING DATE: 1939-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO 4296
: LENGTH: 576
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(576)
: OTHER INFORMATION: unsure at all n locations
: OTHER INFORMATION: Clone ID: L1B3028-013-Q1-B1-G1
:
: US 09-878-574-4296

```

Query Match	8.3%;	Score 76.6;	DB 10;	Length 576;
Best Local Similarity	50.9%;	Pred. No. 3.5e-10;		
Matches 207; Conservative	0;	Mismatches 199;	Indels 1;	Gaps 1

OY	391	CACAGCACTGTACAGATTAATAAAACACACACACCAACCCAAACACACCCGACA-AGCCCACTA	444
Db	556	CACAAAAACCAACACACCCACACAAAAACACACACACAAAAACAAAAACCAACAAA:CAAAA	497
OY	450	CAAAACCAACGCCAAAACCAAAACCCACCCAAAACAAACCCCAATATGATTTTCACTTCGAAGTGT	509
Db	496	ACAACCAACCAACCCACACACACCCACCAAAACACACACAAAAACAAACCAACCA:ACAACCAAAAC	437
OY	510	TTTAACTTTGTACCCCTGCAGCATATGACAGACAATTCACACTGTGGGCTATCTGCAAA	565
Db	436	CCAAAACACCAACCAAAACCCACAAACCAACCAACCAACCAACCAACCAACCC:TAACCAACA	377
OY	570	GAATTCCAAACAAAAAACCCAGAAAGAAAACCCACCACCAAGCTTACAAAAAAACCAACCT	622
Db	376	AACAACCAACACACACACAAAAAAACAAACCAACCCCAACAAACCAACCAACCAACAAAAA	317
OY	630	TCAAGACACACCAAAAAAAGATCTCAACCTCAACCACTAAACCAAGAAAGTAGTACCACCA	689
Db	316	ACACACACACAAAAACCAAAACCCACACAAAAACAAACCAACCAACCAACCAACCCCAACCAAA	257
OY	690	CCAAAGCCACAGAAAGGCAACCATCAACACCAACCAAAACAAACATCAGACATCACTGC	749
Db	256	CAAAAAACACCCCAACACACAAACCAAAACCCACCCACCCCAACCAACCAACCAACACACACA	197
OY	750	TCACCAACACACACACAGGAAATCCAAAACATCCAAATCAATCAATGGA	796
Db	196	CCACCCACACCAAAAAACAAACACCCCAACACACACAAACCAACCAACCAAA:150	

```

RESULT 7
US-09-983-965-2109/c
: Sequence 2109, Application US/09983965
: Patent No. US2002013760A1
: GENERAL INFORMATION:
: APPLICANT: Warren, Wesley C.
: APPLICANT: Tao, Ningbing
: APPLICANT: Bvatt, John C.
: APPLICANT: Mathalagan, Nagappan
: TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
: FILE REFERENCE: 37-21(10297)C
: CURRENT APPLICATION NUMBER: US/09/983, 965
: CURRENT FILING DATE: 2001-10-26
: PRIOR APPLICATION NUMBER: US 09/465, 231
: PRIOR FILING DATE: 1999-12-15
: PRIOR APPLICATION NUMBER: US 60/113, 678
: PRIOR FILING DATE: 1998-12-17
: NUMBER OF SEQ. ID NOS: 5912
: SEQ. ID NO 2109
: LENGTH: 529
: TYPE: DNA
: ORGANISM: Bos taurus
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (335)
: OTHER INFORMATION:

```

OTHER INFORMATION: Clone ID: 34-LIB3057-015-Q1-K1-A6
US-09-983-965-2109

Query Match 8.3%; Score 76.4; DB 10; Length 529;
Best Local Similarity 47.9%; Pred. No. 3.8e-10;
Matches 218; Conservative 0; Mismatches 237; Indels 0; Gaps 0;

```

QY 324 AATTATCATCAAAACCCACCTACTAGTTCATCAACAACCCAGAGTCAAGTCAAC 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 504 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 445
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 384 TCAACCCACAGTCAAGACTAAACACACACACACACACACACACACACACAC 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 444 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 385
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 444 CCAGTCAAAACAGCCCAAAACACACACACACACACACACATATGTTTCCTCG 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 384 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 AAGTGTTCCTTTGACCTGAGCATATGACACACATCCACCTGCTGGCTATCT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 564 GCAGAAATATACCAACCAAAACACAGAAAGAAACACACACACAGCTTACAAAAAC 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 264 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 624 CACCTTCAAGCAACCAAAAGATCTCAACCTCAACCTCAACCAAGGAAGTAC 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 204 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 684 CCACCAACAGCCACAGAGACCAACCATCAACACACACCAAAACATCACAAC 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 144 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAC 85
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 744 CACTGCTCACCAACACACACAGGAATCCAAAA 778
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 84 AACCAACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 50
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

US-09-878-574-4306/c
Sequence 4306, Application US/09878574
Patent No. US20020110548A1
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: Ia Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 4306
LENGTH: 1030
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(1030)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-D11
US-09-878-574-4306

Query Match 8.3%; Score 76.4; DB 10; Length 1030;
Best Local Similarity 41.0%; Pred. No. 5.3e-10;
Matches 285; Conservative 0; Mismatches 399; Indels 11; Gaps 2;
QY 202 AACCCAAAGTCACTAATGCAATCATAAGAGCAACAAGCCAGATCAAGAA 261
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB 793 ACAACACAACACACACACACACACACACACACACACACACACACACACAC 734
QY 262 CACAAACCCCAACATACCTCCTCAGATCCTGATGATGATGATGATGATGATG 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 733 AACCAACACACACACACACACACACACACACACACACACACACACACACAC 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 TGAATTCATCATCAAAACACACACATCTAGTTCACACACACACAGAGTCAAG 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 673 ANA-----MCAACCAACACACACACACACACACACACACACACACACAC 622
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 382 CCTCAACCCCAACACATCAAGACTAAACACACACACACACACACACACACAC 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 621 ACAACCAANNAACACACACACACACACACACACACACACACACACACACAC 562
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 GCCCATACAAACACACACACACACACACACACACACACACACACACACACAC 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 561 NCACACNCANANACACACACACACACACACACACACACACACACACACACAC 502
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 502 CGAAGTGTTCCTTTGACCTGAGCATATGACACACATCCACCTGCTGGGCTAT 561
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 501 ACCCCCCCAGACACACACACACACACACACACACACACACACACACACAC 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 562 CTGCAAAAGATACCAACCAAAACAGAGAAAGAAACACACACACACACACAC 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 441 CACACACACACACACACACACACACACACACACACACACACACACACACAC 382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 622 ACCAACCCTTCAAGCAACCAAAAGATCTCAACCTCAACCTCAACCTCAAC 681
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 AACCAACCAACACACACACACACACACACACACACACACACACACACACAC 325
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 682 ACCCAACCAAGCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 741
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 265
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 742 TACACTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 801
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 264 NACACNACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 205
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 802 CCACTCAACCTCTCTCGAAGCAATCTTAAGCCCTTCTCAAGTCTCCACAC 861
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 204 CNANACNACNACNACNACNACNACNACNACNACNACNACNACNACNACN 145
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 862 CCATCAACACCCCTCATCTCCACCCACACACACAC 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 144 NNANCNANNNNNNNNNNNNNNNNNNNNNNNNNNN 110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 9

US-09-864-761-19241/c
Sequence 19241, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weosheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aesomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30

```

? PRIOR APPLICATION NUMBER: PCT/US01/00667
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00664
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00669
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00665
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00668
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00663
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00662
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00661
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: PCT/US01/00670
? PRIOR FILING DATE: 2001-01-30
? PRIOR APPLICATION NUMBER: US 60/234,687
? PRIOR FILING DATE: 2000-09-21
? PRIOR APPLICATION NUMBER: US 09/608,408
? PRIOR FILING DATE: 2000-06-30
? PRIOR APPLICATION NUMBER: US 09/774,203
? PRIOR FILING DATE: 2001-01-29
? NUMBER OF SEQ ID NOS: 49117
? SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
? SEQ ID NO 19241
? LENGTH: 1075
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? OTHER INFORMATION: MAP TO AL078472.1
? OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
? OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
? OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
? OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
? OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
? OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 43
? OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
? OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
? OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
? OTHER INFORMATION: EST_HUMAN HIT: AV739739.1, EVALDE 1.00e+00
? OTHER INFORMATION: NT HIT: AL163201.2, EVALDE 2.00e-19
US-09-864-761-19241

Query Match      8.3%: Score 76.4; DB 10; Length 1075;
Best Local Similarity 44.5%: Pred. No. 5,5e-10;
Matches 343; Conservative 0; Mismatches 426; Indels 1; Gaps 1;
QY 128 ATCATATTTCATTCTGGCAATGATATCTCACTTATATATACAGCCATCATATA 187
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 909 ACCACACCCACCTACTACCAACACCCACCCACCTACACACCCACCCACCCATCACC 850
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 188 TTCATAGCCTGGCAACCAAGTCAACACTCACTGATGATATATATACAGATCACA 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 849 ACTAAACCATGACCAACCAACCAACCATGACCTACCATCAGCATCAGCATCAGC 790
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 248 ACCGAGATCAAGACACCAACCCCAACATACCTCACTGAGATCTTCAGCTTGGATCAGC 307
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 789 ACCACACCCACCATCAGCAACCAACCAACCATCAGCATCAGCATCAGCATCAGC 730
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 308 TTCTCCAAATCTGTCTGAAATTAATCAACAAACCAACCATAGCTTCAACCAACACA 367
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 729 ACTACACCCACCATCAGCAACCAACCAACCATCAGCATCAGCATCAGCATCAGC 670
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 368 GGAGTCAAGTCAACCTGCAACCCACCAACAGTCAAGCTATTAACCAACCAACCAAC 427
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 669 ACCACACCCACCATCAGCAACCAACCAACCATCAGCATCAGCATCAGCATCAGC 610
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 428 ACACAACCCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAT 487
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 609 ACCACACCCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 550
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```

QY 488 AATGATTTTCACTTGGAGTGTTTACTTTGTACCTTCAGCATATGACGCGAATCA 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 549 ACTACACCCACCATCAGCAACCAACCAACCATCAGCATTAACCAACCAACCAACCAAC 490
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 548 ACCTGCTGGGTATCTGCAAAAGAAATACCAACCAACCAACCAACCAACCAACCAAC 607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 489 ACCATCATCAGCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 430
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 608 AAGCTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 429 ACCACCATCAGCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 370
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 668 AAACCAAGGAAGTATCCGACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 369 ACCATCATCAGCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 310
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 728 ACAACATCAACATCACTGCTGACCAACCAACCAACCAACCAACCAACCAACCAAC 787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 309 ACCACCATCAGCATCAGCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 250
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 788 CAATGGAACCTTCTCACTCAACCTCTCGGAGGCA-ATCTAAGCCCTTCTCAAGTCTC 846
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 249 ACCACCATCAGCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 847 CACAACATCGAGCAGCAGCATCAACCAACCAACCAACCAACCAACCAACCAAC 896
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 189 ACCACCATCAGCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
US-09-864-761-2513/c
: Sequence 2513, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Aeomica X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
```

PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 2513
LENGTH: 1403
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL078472.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
US-09-864-761-2513

Query Match 8.3%; Score 76.4; DB 10; Length 1403;
Best Local Similarity 44.5%; Pred. No. 6.2e-10;

Matches 343; Conservative 0; Mismatches 426; Indels 1; Gaps 1;

128 ATCAATTATTCATTTGCGATGATATCTCACTTCACTTAAATTAAGCATCATA 187
1259 ACCACCCACCACTACACCAACACACACACACACACACACACACACACACAC 1200
188 TTCAATGCTTGGCAACCAAGTCACTAATCACTGCAATCATACAGATGACA 247
1199 ACTAAACCATGACACACACACACACACACACACACACACACACACACAC 1140
248 AGCAGATCAAGAACACACACACACACACACACACACACACACACACACAC 307
1139 ACCACCCACCACTACACACACACACACACACACACACACACACACACAC 1080
308 TTCTCAATGCTGTAATTAATGATCAACACACACACACACACACACACAC 367
1079 ACTAC 1020
368 GGAATGAGTCAACCTGCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 427
1019 ACCACCCACCACTACACACACACACACACACACACACACACACACACAC 960
428 ACACACCCACCACTACACACACACACACACACACACACACACACACACAC 487
959 ACCACCCACCACTACACACACACACACACACACACACACACACACACAC 900
488 AATGATTTGCTGGAAGTGAATTTGATCTGACGATATGACGACCAATCA 547
899 ACTAC 840
548 ACCGCTGGGCTATCTCAAAAGATACAAAGAAAAACAGAAAAAGAAAAAC 607
839 ACATCTAC 780
608 AAGCTTCAAAAAACCACTTCAAGACCAAAAAAAGATCTCAAACTCAACCACT 667
779 ACCACCCACCACTACACACACACACACACACACACACACACACACACAC 720
668 AAACCAAGAGAGTACCAACCAACCAACCAACCAACCAACCAACCAACCA 727
719 ACACACTACATCAACCACTACACACACACACACACACACACACACACAC 660
728 ACAACATCACTACATGCTGCTACCAACCAACCAACCAACCAACCAACCA 787

Db 659 ACCACCATCACCACCTACCATCACCACACACACACACACACACATCACCAC 600
Qy 768 CAATGGAACCTTCCACTCAACCTCCGCGAAGGCA-ATCTAGCCCTTCAAGTCTC 846
Db 599 ACCACCATCACCACACACACACACACACACACACACACACACACACACATCACC 540
Qy 847 CACAATCCGAGACACACACACACACACACACACACACACACACACACAC 896
Db 539 ACCACCATCACCACACACACACACACACACACACACACACACACACACATCACC 490

RESULT 11
US-09-878-574-4312/c
Sequence 4312, Application US/09878574
Patent No. US20020110548A1

GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 4312
LENGTH: 577
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)-(577)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-G11
US-09-878-574-4312

Query Match 8.2%; Score 75; DB 10; Length 577;
Best Local Similarity 50.0%; Pred. No. 9.2e-10;
Matches 232; Conservative 0; Mismatches 230; Indels 2; Gaps 2;

Qy 336 AAACACACACATCTACTAGCTTCAACACACACACACACACACACACACAC 395
Db 533 AAAC 474
Qy 396 GAGTAAAGCTAAACACACACACACACACACACACACACACACACACACAC 455
Db 473 AAAC 414
Qy 456 AAAC 514
Db 413 CACAAC 354
Qy 515 TTGTACCTTGAGATATGACGACAAACATCCAACTGCTGGGTATGTGAAAGATA 574
Db 353 CAAAAAACCCACACACACACACACACACACACACACACACACACACAC 294
Qy 575 CAAAAAACCCACACACACACACACACACACACACACACACACACACAC 633
Db 293 CCAAC 234
Qy 634 GACAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCAAAAGAGTACCAACCA 693
Db 233 AAAAAAACCCACACACACACACACACACACACACACACACACACACAC 174
Qy 694 GCCACACAAAGGCAACACATCAACACACACACACACACACACACACAC 753
Db 173 CACCCCAAAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCAAAACCA 114
Qy 754 CAAAC 797
Db 113 CAAAC 70

```
RESULT 12
US-09-878-574-4316/c
: Sequence 4316, Application US/09878574
: Patent No. US20020110548A1
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/333,535
: PRIOR FILING DATE: 1999-06-14
: NUMBER OF SEQ ID NOS: 15775
: SEQ ID NO 4316
: LENGTH: 639
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: unsure
: LOCATION: (1)..(639)
: OTHER INFORMATION: unsure at all n locations
: OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-H11
US-09-878-574-4316

Query Match      8.08; Score 74; DB 10; Length 639;
Best Local Similarity 49.18; Pred. No. 1,8e-09;
Matches 223; Conservative 0; Mismatches 230; Indels 1; Gaps 1;

QY 324 AAATTACATCAACAACCCATCTAGCTTCAACACACGAGATCAAGTCAACC 383
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 AAACAACACAAACACAAACAAACAAACAAACACACACACACACAAACCA 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 384 TGCAACCCACACAGTCAAGCTAAAAACACACACACACACACACACACAC 443
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 506 ACAACACAAACACACACACACACACACACACACACACACACACACACAA 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 444 CCATACACAAACACACACACACACACACACACACACACACACACACACAC 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 446 ACAACACACACACACACACACACACACACACACACACACACACACACAC 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 AAGTGTTCATCTGTACCTGTCAGCATATGACACACATCCAACTGCTGGCTATCT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 AACACACACACACACACACACACACACACACACACACACACACACACAA 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 564 GCAAAAGCATACCAAAACAAACACAGGAAAGAAACACACACACACACACAC 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 CAACACACAAACAAACAAACAAACACACACACACACACACACACACACAC 268
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 624 CAACCTTCACACACACACACACACACACACACACACACACACACACACAC 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 ACACAAACACACACACACACACACACACACACACACACACACACACACAA 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 684 CCACACACACACACACACACACACACACACACACACACACACACACACAC 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 207 ACACACACACACACACACACACACACACACACACACACACACACACACAC 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 744 CACTGCTCACCACACACACACACACACACACACACACACACACACACAC 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 147 CAACACACACACACACACACACACACACACACACACACACACACACAA 114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 13
US-09-864-761-26582/c
: Sequence 26582, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
```

```
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecomics-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 26582
: LENGTH: 576
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AP000111.1
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.2
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
: OTHER INFORMATION: SWISSPROT HIT: P44836, EVALUUE 1.00e-12
US-09-864-761-26582

Query Match      7.68; Score 70; DB 10; Length 576;
Best Local Similarity 56.7%; Pred. No. 1,8e-08;
Matches 170; Conservative 0; Mismatches 125; Indels 5; Gaps 2;

QY 427 AACACAAACCCAGACAGCCCTACAAACACAGCCCAAAACCAACCAACAA 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 569 AAGAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 487 TAATGATTTTCACTGGAAGTGTAACTTGTACCTGACAGATATGACGAACATCC 546
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 509 CAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 453
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 547 AACCTGTGGCTATCTGCAAAAGAAATACCAAAACAAACCAAGAAAGAAACCA 606
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 452 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 607 CAAGCTACAAACCAACCAACCTTCAAGACACCAACCAACCAACCAACCA 666
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db 1350 AAAACATTAACACATATAAAAAATTAATAAAAAATAAACACCAAAAAATA 1409
OY 803 CAC 805
Db 1410 AAC 1412

Search completed: May 1, 2003, 13:43:09
Job time : 108.532 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:08:32 ; Search time 1137.76 Seconds
(without alignments)
13095.757 Million cell updates/sec

Title: US-09-462-816-1
Perfect score: 920
Sequence: 1 tgcacacatgccaaaaa.....gtatattataaaaaaa 920

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estda:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120	13.0	895	17	CNS0071A
2	104.6	11.4	946	17	AG127412
3	103.2	11.2	922	17	CNS0073W
4	101.8	11.1	919	17	CNS04ENV
5	100	10.9	1201	9	AL581589
6	99.4	10.8	1141	17	AQ743305

C	7	99	10.8	1101	17	CNS00FXE	AL071370	Drosophila
C	8	96.6	10.5	1024	17	AG133080	AG133080	Pan trogl
C	9	95.8	10.4	1097	17	CNS01224	A1102202	Drosophila
C	10	95.8	10.4	1101	17	CNS0100X	A1038379	Drosophila
C	11	95	10.3	791	17	CNS009KS	AL053801	Drosophila
C	12	95	10.3	1101	17	CNS00LOO	AL053807	Drosophila
C	13	95	10.3	1144	17	AG743364	AG743364	HS_5387_B
C	14	94.4	10.3	1216	17	AG135357	AG135357	Pan trogl
C	15	94.2	10.2	815	17	B12686	B12686	F27J14-77 I
C	16	94	10.2	1188	17	AG135532	AG135532	Pan trogl
C	17	93.8	10.2	859	17	AG128925	AG128925	Pan trogl
C	18	93.8	10.2	1008	17	AG137085	AG137085	Pan trogl
C	19	93.8	10.2	1099	17	AG743360	AG743360	HS_5387_B
C	20	93.6	10.2	884	17	CNS006U0	AL065923	Drosophila
C	21	93.6	10.2	969	17	AG743309	AG743309	HS_5387_B
C	22	93.6	10.2	1081	17	AG135328	AG135328	Pan trogl
C	23	92.8	10.1	955	17	AG076494	AG076494	Pan trogl
C	24	92.8	10.1	981	17	AG127518	AG127518	Pan trogl
C	25	92.8	10.1	1353	17	AG128010	AG128010	Pan trogl
C	26	92.6	10.1	1183	17	AG136828	AG136828	Pan trogl
C	27	92.4	10.0	880	17	AG139490	AG139490	Pan trogl
C	28	92.2	10.0	885	17	CNS031TM	AL249807	Tetradon
C	29	92	10.0	1101	17	CNS00LT2	AL078714	Drosophila
C	30	92	10.0	1137	17	AQ743326	AQ743326	HS_5387_B
C	31	92	10.0	1225	17	CNS0166K	AL105358	Drosophila
C	32	92	10.0	1491	12	BE882936	BE882936	601505836
C	33	91.4	9.9	700	17	AG127423	AG127423	Pan trogl
C	34	91.4	9.9	866	17	AG126308	AG126308	Pan trogl
C	35	91.2	9.9	1054	17	AG137135	AG137135	Pan trogl
C	36	91	9.9	1749	13	BM413579	BM413579	Op20453 M
C	37	90.6	9.8	927	17	AG139684	AG139684	Pan trogl
C	38	90.6	9.8	987	17	CNS0142W	AL103634	Drosophila
C	39	90.2	9.8	1026	17	AG126150	AG126150	Pan trogl
C	40	90	9.8	953	17	CNS006PF	AL056567	Drosophila
C	41	90	9.8	1101	17	CNS017FC	AL107970	Drosophila
C	42	89.8	9.8	896	17	AG141027	AG141027	Pan trogl
C	43	89.8	9.8	1102	17	AG135459	AG135459	Pan trogl
C	44	89.4	9.7	909	17	CNS00JTL	AL076720	Drosophila
C	45	89.4	9.7	921	17	AG127701	AG127701	Pan trogl

ALIGNMENTS

RESULT 1
CNS0071A/c 895 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACRI4B09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL066286
VERSION AL066286.1 GI:4945153

KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 895)
Genoscope.

AUTHORS Direct Submision
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BR 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial

OY	820	AGGGAATCTAAGCCCTTCTCAAGTCCTCACAACACTCGAGACGCCATCATCACACCCTCATC	879
Db	753	CAGCAGCCCCCAAAAAAAAAAACCCCCCACCACACAAAAAAACAGACACAAACACACCACAA	812
OY	880	TTCACCCACACACAGCCAGTGATTATTTAAAAAAA	919
Db	813	CCAAAACCCCACAGAAAAACGACACAGCCCCACACACAAAAA	852
 RESULT 3 CNS0073M/c			
DEFINITION Drosophila melanogaster genome survey sequence TET3 end of BAC #.			
LOCUS BACR14D09 of RPCI-98 library from Drosophila melanogaster (fruit-			
fly), genomic survey sequence.			
AL066784			
AL066784.1 GI:4945247			
GSS.			
Drosophila melanogaster.			
Drosophila melanogaster.			
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
Ephydroidea; Drosophilidae; Drosophila.			
1 (bases 1 to 922)			
Genoscope.			
Direct Submission			
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;			
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr)			
Web : www.genoscope.cns.fr			
Determination of this BAC-end sequence was carried out as part of a			
collaboration with the Berkeley Drosophila Genome Project (BDGP).			
The BDGP is constructing a physical map of the Drosophila			
melanogaster genome using these BACS. For further information			
please see http://www.fruitfly.org The BDGP Drosophila			
melanogaster BAC library was prepared by Kazutoyo Osoegawa and			
Aaron Mammosser in Pieter de Jong's laboratory in the Department of			
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,			
NY. The library is named RPCI-98 and was constructed by partial			
EcoRI digestion of Drosophila DNA provided by the BDGP from the			
isogenic strain y2; cn bw sp, the same strain used for the BDGP's			
P1 and EST libraries. A more detailed description of the library			
and how to order individual BAC clones, the entire library, or			
filters for hybridization from the BACPAC Resource Center can be			
found at location://bacpac.med.buffalo.edu/drosophila_bac.htm.			
Location/Qualifiers			
1..922			
/organism="Drosophila melanogaster"			
/db_xref="taxon:7227"			
/clone_lib="BACR14D09"			
/note="end : TEF3"			
BASE COUNT 223 a 95 c 109 g 221 t 274 others			
ORIGIN			
Query Match 11.2%; Score 103.2; DB 17; Length 922;			
Best Local Similarity 24.8%; Pred.No.1.9e+10;			
Matches 102; Conservative 164; Mismatches 144; Indels 1; Gaps 1;			
OY	386	CAACCCACACAGTGAAGAATAAAACACACACACACACACACCCACACGGCAGCC	445
Db	921	MMMCMCMCCMCCCCCMAC	862
OY	446	ACTCAAAAACAGGCAACAAACACCAACCAACCAACCAATTAATTTCACTTGC	505
Db	861	ACAMMAAMNMNMACMNMNMACMNMNMACMNMNMACMNMNMACMNMNMACMNMNMAC	802
OY	506	GTTTAACTTTGTACCTCGACGATATGACGAACAATCAACCTGCTGGGTATCTGC	565
Db	801	ACACMCAMNMCMNMNMNMNMNMNMCMCA-CNMNACMNMCMCMCMCMCMCMCMNA	743
OY	566	AAAAGATTCGCAAAAAACAGGAAGAAACCAACACACACACAGCTTAGAAAAACA	625

[illegible]

BASE COUNT	92 a	57 c	490 g	461 t	101 others
ORIGIN					
Query Match		10.98;	Score 100;	DB 9;	Length 1201
Best Local Similarity		45.5%;	Pred. No. 8e-10;		

	Matches	256; Conservative	39; Mismatches	264; Indels	4; Gaps	2
QY	337	AAACACCCACCTACTACTAGTTCTCAACAACACCCAGAGTCAAGTCAACCTCTCAACCCACAC	396			
Db	1152	AA	1093			
QY	397	AGTCAAGACTAAAGAACACACAAACACACCCCAACACACACCCACAGCCCTACTACAAACAA	456			
Db	1092	AACCCNCAAAC	1033			
QY	457	ACGCCAAAAAACACCAACCAACCCCAATATGATTTTCACTTGGATTTTAACTT	516			
Db	1032	ACCCCAACAAAAAACCAACCCCAACAAAAAAMCCCMNACAAAAAAMAAAAAMANCA	973			
QY	517	TGTACCTGACACTATGTGACGACAAATCCAACTCTGTGGCTATCTGTGAAAAAGATACC	576			
Db	972	AAACACAAACCCCAAAAAAAAAAACCCAAACCCCAACAAAAAACAAACAAACAAACCMCAAAAC	913			
QY	577	AAACAAAAAACCCAGAAAGAAAAACCAACCAACCAAGCTTACAAAAAACCAACTCTTAAGAC	636			
Db	912	MMCCMAACAAACMAAAAAAAAAAACCA--AACAAACCAAAAAAAAAAAAAAAAAACCAACMAA	855			
QY	637	AAACAAAAAAGATCTCAAACTCAAAACCTAAACCAAGGAAGTACCACCAACCAAGCC	696			
Db	854	AAC--AACACACCCCAACAAAAAAMAAAAAAMAAAAAACAAAAAACACAAACAAAAAACAMA	797			
QY	697	CACGAGAGGCCAACCATCAACACGCCAAACAAACAAACATCACTACTACAGTCTCAACAA	756			
Db	796	AAAAAAAAAAMCACAACAAAAAACAAAAAAMAAAAACMAAACCAACCAACCAAAAAAMAAA	737			
QY	757	CACACACACAGAGAAATCCAAAAATCAAGTCAATATGAGAAACCTTCACATCAACCTCTGC	816			
Db	736	MAAACACCCCAAMMCAAAAAAAAAAACAAAAAACCMMAAACCAACCCCAACCAACCCCAAC	677			
QY	817	CGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACAATCTCGAGACCCATCAACAACCTC	876			
Db	676	CACCCCCCAACACACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCC	617			
QY	877	ATCTCAGCCCAACACACACAGCC	899			
Db	616	CCCMCCCCCCCCCCCCCCCCC	594			

	RESULT	6	
A0743305	Locus		
A0743305	Definition	1141 bp	DNA linear GSS 16-JUL-1999
		HS_5387_B2_B02.SP6 RPlCI-II Human Male BAC Library Homo sapiens genomic clone Plate=963 Col=4 Row=D, DNA sequence.	
	Accession	A0743305	
	Version	A0743305.1	GI:5520827
	Keywords	GSS.	
SOURCE	Organism	human.	
	Homo sapiens		
	Eumariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE		1 (bases 1 to 1141)	
AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J.J., Young,J., Zhao,S., Adams,M.D. and Hood,L.		
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome		
JOURNAL MEDLINE COMMENT	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999) 99380589		
Contact:	Mhairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel.: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu		
	BAC library availability, please contact Pieter de Jong (pietred@jg.med.buffalo.edu). Clones may be purchased from		

BACPAC Resources (http://bacpac.med.bufileo.edu/ordering_bac.htm)
or from Research Genetics (<http://www.htsc.washington.edu>). BAC end Web Server:
<http://www.htsc.washington.edu>
Plate: 963 row: D column: 4
Seq primer: SP6
Class: BAC ends

High quality sequence stop: 1141.

FEATURES

source

1. 1141

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="Plate:963 Col-4 Row-D"

/clone_11b="RPCT-11 Human Male BAC Library"

/sex="male"

/note="Vector: PBACE3.6; Site 1: EcoRI; Site 2: EcoRI;

Male blood DNA was isolated from one randomly chosen donor

and partially digested with a combination of EcoRI and

EcoRI Methylase. Size selected DNA was cloned into the

PBACE3.6 vector at EcoRI sites"

BASE COUNT

760 a 301 c 16 g 15 t 49 others

ORIGIN

Query Match

Best Local Similarity 46.7%; Score 99.4; DB 17; Length 1141;

Matches 374; Conservative 0; Mismatches 417; Indels 10; Gaps 4;

120

14

180

74

240

134

300

194

360

248

420

306

480

366

540

426

600

486

660

545

720

605

780

665

DB

DB

DB

DB

DB

DB

DB

DB

QY 840 AAGTCCACACACATCCGACCCATGCAACCTCATCTGCACCCACACACACGCC 899
DB 725 AAAAC-CAAAAACACACATCCACACCCACAAAACACACAAAAAACAACACAC 783
QY 900 ACTAGTTATTTAAAAA 920
DB 784 ACAACACAAACCAAAAA 804

RESULT 7

CNS00FHE/C

LOCUS

DEFINITION

BACR32C19 of RPCT-98 library from Drosophila melanogaster (fruit

fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster.

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1101)

Genoscope.

Direct Submission

Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage ;

Bp 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP).

The BDGP is constructing a physical map of the Drosophila

melanogaster genome using these BACs. For further information

please see <http://www.fruitfly.org> The BDGP Drosophila

melanogaster BAC library was prepared by Kazutoyo Osoegawa and

Aaron Mamoser in Pletier de Jong's laboratory in the department of

Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,

NY. The library is named RPCT-98 and was constructed by partial

EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's

P1 and EST libraries. A more detailed description of the library

and how to order individual BAC clones, the entire library, or

filters for hybridization from the BACPAC Resource Center can be

found at http://bacpac.med.bufileo.edu/drosophila_bac.htm.

FEATURES

Location/Qualifiers

1. 1101

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/clone="BACR32C19"

/clone_11b="RPCT-98"

/note="end : TET3"

BASE COUNT

326 a 104 c 183 g 186 t 302 others

ORIGIN

Query Match

Best Local Similarity 25.6%; Score 99; DB 17; Length 1101;

Matches 104; Conservative 145; Mismatches 157; Indels 0; Gaps 0;

358

DB

QY

DB

QY

DB

QY

DB

QY 358 AACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 417
DB 1081 AAACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 1022
QY 418 AACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 477
DB 1021 AAACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 962
QY 478 AACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 537
DB 961 AAACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 902
QY 538 AACACACGAGCTCAAGTCAAACTGCAACCCACACAGCTCAAGCTAAACACACAC 597

QY	598	155	202	215	262
QY	598	155	202	215	262
Db	841	155	202	215	262
QY	658	155	202	215	262
Db	781	155	202	215	262
QY	718	155	202	215	262
Db	721	155	202	215	262
RESULT 8					
LOCUS	AG133080				
DEFINITION	Pan troglodytes DNA, clone: PTB-145K08.R, genomic survey sequence.				
ACCESSION	AG133080				
VERSION	AG133080.1				
KEYWORDS	GSS.				
SOURCE	Pan troglodytes male lymphoblast DNA, clone_11b:PTB Chimpanzee Male BAC library clone:PTB-145K08.R.				
ORGANISM	Pan troglodytes				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.				
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	BAC end sequences of library PTB				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 1024)				
AUTHORS	Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Tokoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: chimbesegsc.riken.go.jp, URL: http://bgp-gsc.riken.go.jp/); Tel:81-45-503-9111, Fax:81-45-503-9170)				
COMMENT	Clones are derived from the chimpanzee BAC library PTB This BAC end was generated during the Rdb process and may have higher chance of clone tracking errors.				
PRIMERS					
LIBRARY	Sequencing: MJRev				
VECTOR	: pRS145				
R SITE 1	: SacI				
R SITE 2	: SacI.				
LOCATION/QUALIFIERS	1..1024				
ORGANISM	"Pan troglodytes"				
DB_XREF	"taxon:9598"				
CLONE	"PTB-145K08.R"				
SEX	"male"				
CELL_TYPE	"lymphoblast"				
CLONE_11b	"PTB Chimpanzee Male BAC library"				
BASE COUNT	495 a 460 c 23 g 40 t 6 others				
ORIGIN					
Query Match	10.5%; Score 96.6; DB 17; Length 1024;				
Best Local Similarity	47.3%; Pred. No. 3.7e-09;				
Matches	354; Conservative 0; Mismatches 390; Indels 4; Gaps 2;				
QY	155	ATCTCACTTCACTTATTAATACAGGCAATCATTTATGCTCGGCAAAACCAAGTC	214		
Db	202	AACAAACCCCAACACACCAACCAACCAACCCCACTCCCAACAAACCAAAAC	261		
QY	215	ACACAAACACGCAATCATTAAGATGCAACAAAGCAACAAACCAACCA	274		
Db	262	AACCCCAACAAACCCCAACCAATTAACAAACACACCAACCAACCAACCAACCC	321		

[illegible]

LOCUS	CNS009KS	791 bp	DNA	linear	GSS 03-JUN-1996
DEFINITION	Drosophila melanogaster genome survey sequence T7 end of BAC # BRG19F04 of Rpcl-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL053801				
VERSION	AL053801.1 GI:4935176				
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster.				
ORGANISM	Drosophila melanogaster. Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 791) Genoscope.				
AUTHORS	Direct Submission				
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)				
JOURNAL					
COMMENT	<p>- Web : www.genoscope.cns.fr) determination of this BAC end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazuo Osoegawa and Aaron Mamoser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named Rpcl-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain Y2: cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.</p>				
FEATURES	Location/Qualifiers				
source	1..791 /organism="Drosophila melanogaster" /db_xref="taxon:7227" /clone="BACR19F04" /clone_1lb="Rpcl-98" /note="end : 17"				
BASE COUNT	464 a 105 c 67 g 78 t 77 others				
ORIGIN					
Query Match	10.3%; Score 95; DB 17; Length 791;				
Best Local Similarity	47.1%; Pred. No. 7.7e-09;				
Matches	205; Conservative 43; Mismatches 178; Indels 9; Gaps 2.				
OY	358 AACCAACCAGGAGTCGAAGTCCAACTGCAGACCCCACAAGTCAGACTCAAACGACANC	417	:	:	:
Db	33 AAAAAAAAAAAMCMCAAAAAAACACAAAATAAAAAAAAAAACGGAAGAAAAACAA	92	:	:	:
OY	418 AACCAACCACAAACCCAGCAGGCCACTACAACAAACAGCCCAAAACCAACCAANA	477	:	:	:
Db	93 AAAAAAAAAAAMCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA	152	:	:	:
OY	478 CAATCCCA-ATTATGTCTTCACTTGCGAAGTTTAACCTTGATCCCTGCACATATGCA	536	:	:	:
Db	153 AAAAAAAAAAAMMAAAMMAAAMMAAACAACAAACACAAAAAAMAGAACAACAGGA	212	:	:	:
OY	537 GCACAAATCCAACCTGGTGGCTATCTGCMAAAGATATACCAACAAACCAAGGAAGA	596	:	:	:
Db	213 ACWATTAAGCACAAA-----AAAMMMAAAMMAAAMMAAAMMAAAMMAAAMMAA	264	:	:	:
OY	597 AAACCAACCAAGCCTACAAAAAACCAACCTTCAAGCAACCAACCAAAAAAGATCTCAAC	656	:	:	:
Db	265 AAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	324	:	:	:
OY	657 CTCACCAACCACTAAACCAAGGAAGTACCAACCAAGCCACAGAGAGGCAACCATCA	716	:	:	:
Db	325 AAAAAAAAAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAAAMMAA	384	:	:	:
OY	717 ACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	776	:	:	:

[illegible]

```

Db      695  MTTAATTCACIMMAMAMMTTMTTMMAMACSRMAMMMCCSCSAMCMCAMMAMMMACMA 754
Oy      609  AGGCTTACAAAAMAAACACCTTCAAGACACCAAAAAGATCTCAAACTCAACCACTA 668
Db      755  WIMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 814
Oy      669  AACCAAGAGTACCCACCCACCAAGCCACAGAGAGCCACATCAACCCACCAAAA 728
Db      815  YACSCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 874
Oy      729  CAACATTCACACTACCTGTCTACCAACACACACAGCAATTCACCAAGTCTCAAGT 788
Db      875  CCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 934
Oy      789  AATGGAACCTTCCACATCCATCCGAGGACATCTAAGCCCTTCAAGTCTCA 848
Db      935  MAMMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 994
Oy      849  CAACATCCGAGCACCACATCAACACCTTCATCTCCACCCACACACACAGCCAGTA 903
Db      995  MAMMAMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCMCM 1049

```

```

RESULT 13
LOCUS   A0743364 1144 bp DNA linear GSS 16-JUL-1999
DEFINITION HS.5387_B2_H09_SP6 RPCI-11 Human Male BAC Library Homo sapiens
ACCESSION A0743364
VERSION   A0743364.1 GI:5520886
KEYWORDS  GSS.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
           1 (bases 1 to 1144)
           Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
           Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
           Hood,L.
           Sequence-tagged connectors: A sequence approach to mapping and
           scanning the human genome
           Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
           99380589

```

```

TITLE    Contact: Mahairas GG, Wallace JC, Hood L
JOURNAL   High Throughput Sequencing Center
MEDLINE   University of Washington
COMMENT   401 Queen Anne Avenue North, Seattle, WA 98109, USA
           Tel: (206) 616-3618
           Fax: (206) 616-3887
           Email: jwallace@u.washington.edu

```

```

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieret@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.hsc.washington.edu
Plate: 963 row: P column: 18
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 1144.
Location/Qualifiers
1..1144
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_1lib="Plate=963 Col=18 Row=P"
/clone_1lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACE3.6 vector at EcoRI sites"

```

```

BASE COUNT 625 a 443 c 11 g 11 t 54 others
ORIGIN

```

```

Query Match 10.3%; Score 95; DB 17; Length 1144;
Best Local Similarity 46.7%; Pred No. 7.5e-09;
Matches 371; Conservative 0; Mismatches 420; Indels 3; Gaps 3;

```

```

Oy      125  CAATTCACATTTATTCATTTCTGGCAATGATTAATCTCACTTCACTTATTAATCAGGCATC 184
Db      340  CAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 399
Oy      185  ATTTGATGATGCTTGGGCAACGCAAAAGTCACACTTAACACTGCAATCTTACAAATGCA 244
Db      400  AAAAANAANAACACACACACACACACACACACACACACACACACACACACACACAC 459
Oy      245  ACAAGCAGATCAAGTCAAGACACACACACACACACACACACACACACACACACACAC 303
Db      460  CCACACACACACACACACACACACACACACACACACACACACACACACACACACAC 519
Oy      304  CAGCTTCTGCAATCTGTCTGAAATTTACATCACAACACACACACACACACACACACAC 363
Db      520  CCACACACACACACACACACACACACACACACACACACACACACACACACACACAC 579
Oy      364  ACCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 423
Db      580  CCACACACACACACACACACACACACACACACACACACACACACACACACACACAC 638
Oy      424  CCAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 483
Db      639  ACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 698
Oy      484  CAATATGATTTTCACTTCTGAAAGTGTACCTTTGACCTTGTACCTTGTACCTTGTAC 543
Db      699  CAACACACACACACACACACACACACACACACACACACACACACACACACACACAC 757
Oy      544  TCCACCTGCTGGCTATCTGCAAAAGATACCAACCAACCAACCAACCAACCAACCAAC 603
Db      758  CAAAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 817
Oy      604  CACCAAGCTTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 663
Db      818  AACCCACACACACACACACACACACACACACACACACACACACACACACACACACAC 877
Oy      664  CACTAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 723
Db      878  ACACACACACACACACACACACACACACACACACACACACACACACACACACACAC 937
Oy      724  CAAAACACACACACACACACACACACACACACACACACACACACACACACACACAC 783
Db      938  CACCCACACACACACACACACACACACACACACACACACACACACACACACACACAC 997
Oy      784  AAGTCAATGGAACCTTCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCA 843
Db      998  CACCCACACACACACACACACACACACACACACACACACACACACACACACACACAC 1057
Oy      844  CTCACACACACACACACACACACACACACACACACACACACACACACACACACACAC 903
Db      1058  AAANAANAACACACACACACACACACACACACACACACACACACACACACACACAC 1117
Oy      904  GTTATTTAAAAAAA 917
Db      1118  CCAACACACACACACACACACACACACACACACACACACACACACACACACACAC 1131

```

```

RESULT 14
LOCUS   AG135357 1216 bp DNA linear GSS 04-NOV-2001
DEFINITION Pan troglodytes DNA, clone: PTB-148j19.F, genomic survey sequence.
ACCESSION AG135357
VERSION   AG135357.1 GI:16665035
KEYWORDS  GSS.
SOURCE    Pan troglodytes male lymphoblast DNA, clone_1lib:PTB Chimpanzee Male
           BAC library clone:PTB-148j19.F.

```

ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Pan.

REFERENCE 1
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Toto, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of library PTB
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1216)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,
Toto, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail: chimbes@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/
Tel: 81-45-503-9111, Fax: 81-45-503-9170)
COMMENT Clones are derived from the chimpanzee BAC library PTB. This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI
Location/Qualifiers
1. 1216
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="PTB-148J19.F"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"
BASE COUNT 572 a 532 c 15 g 29 t 68 others
ORIGIN

Query Match 10.3%; Score 94.4; DB 17; Length 1216;
Best Local Similarity 47.1%; Pred. No. 9,9e-09;
Matches 319; Conservative 0; Mismatches 353; Indels 6; Gaps 2;

222 CAACTGCAATCATACAGATGACAGACAGAGATCAAGCAACCCCACTACCTCA 281
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 109 CAAAANNANACCNCAANCAACAAACCAAAAAACANNAANCCCTCCANCCCAAN 168
282 CTCGAGATCTCAGCTTGAATCAGCTTCTCAATCTGTCTGAATTTACATCAACA 341
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 169 AANAANCCACACCAAAATTCACCCNCAACAAACCCCAACCCCAACCCCAACA 228
342 CCACCACTAGCTTCAACACACACAGAGATCAAGTCAACCTGCAACCCCAACAGTCA 401
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 229 AACCCGACACACACACACCCCAACCAACCCCAACCCCAACCCCAACCAACA 288
402 AGACTAATAACACACACACACCAACCAACCCCAACCCCACTACAAACAAGCGC 461
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 289 CACCCNCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 348
462 AANAACCAACCAACCAACCAACCAATATGATTTTCACTTCAAGTGTAACTTTGAC 521
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 349 CCACCAAAACACCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 403
522 CCGTGAAGATATGACAGCAACATCACTGCTGGGCTATCTGCAAAAGATATCAACA 581
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 404 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 463
582 AAAAAACCAAGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 641
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 464 CACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 523
642 AAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCA 701
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 524 AAAAAAACCAACCAATCTCTGCTACCAACCAACCAACCAACCAACCAACCAACCA 583

QY 702 AAGAGCCAAACATCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 761
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 584 CACAAACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 642
762 CCACAGAAATCCAAAACCTACAAATGAAACCTTCCATCACTCTCCGAG 821
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 643 CCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 702
QY 822 GCAATCTAAGCCCTTCTCAAGTCTCACAACATCCGAGCACCATCAACCTCATCTC 881
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 703 CCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 762
QY 882 CACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 899
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 763 CACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 780

RESULT 15
B12686/c 815 bp DNA linear GSS 14-MAY-1997
LOCUS B12686/c
DEFINITION F27J14-T7 IGF Arabidopsis thaliana genomic clone F27J14, DNA
sequence.
ACCESSION B12686
VERSION B12686.1 GI:2093806
KEYWORDS GSS.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eustroids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 815)
Feng, J., Dewar, K., Buehler, E., Kim, C., Li, Y., Shinn, P., Sun, H. and
Ecker, J.
BAC End Sequences at ATGC
JOURNAL Unpublished (1997)
CONTACT Ecker, J
Arabidopsis Thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@atgenome.bio.upenn.edu
Seq primer: T7
Class: BAC ends
High quality sequence start: 400
High quality sequence stop: 417.
Location/Qualifiers
1. 815
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="F27J14"
/clone_lib="IGF"
/sex="hemaphrodite"
/note="vector: BelobacII; Site_1: EcoRI; Site_2: EcoRI;
Produced by Thomas Altmann"
BASE COUNT 22 a 11 c 185 g 502 t 95 others
ORIGIN

Query Match 10.2%; Score 94.2; DB 17; Length 815;
Best Local Similarity 45.0%; Pred. No. 1.1e-08;
Matches 297; Conservative 0; Mismatches 355; Indels 8; Gaps 1;

220 AACCACTCAATCATACAGATGCAACAGCAGATCAAGACACCAACCCCAACATACCT 279
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 812 AANAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 753
280 CACTAGATCTCTCAGCTTGAATCAGCTTCTCAATCTGTCTGAATTTACATCAAC 339
111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
Db 752 AACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693
340 CACCACTACTAGCTTCAACACACAGAGATCAAGTCAAACTGCAACCCCAACAGT 399

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 12:33:38 ; Search time 37.6717 Seconds
(Without alignments)
1054.072 Million cell updates/sec

Title: US-09-462-816-2
Perfect score: 1544
Sequence: 1 MSKNKDQRTAKLEKTWDTL.....VSTSEHPSPSSPPTTRQ 298

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A_Geneseq_101002:*
2: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/genescp-emb1/AA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/genescp-emb1/AA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/genescp-emb1/AA2001.DAT:*
24: /SIDS2/gcgdata/geneseq/genescp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1544	100.0	298	14 AAR39286	Respiratory syncyt
2	1544	100.0	298	20 AAW96313	Membrane bound G p
3	1418	91.8	298	8 AAP70845	Sequence of human
4	1418	91.8	298	13 AAR25302	HSRV glycoprotein
5	1418	91.8	298	19 AAW47605	HSRV glycoprotein
6	1418	91.8	298	23 AAU74676	Respiratory syncyt
7	1229	79.6	232	20 AAW96314	Secreted G protein
8	901	58.4	681	10 AAP90441	Chimeric human res
9	771	49.9	299	22 AAB68336	Amino acid sequenc
10	572.5	37.1	452	22 AAB67771	Amino acid sequenc

11	562	36.4	349	17	AAR95660
12	562	36.4	349	22	AAB68028
13	558	36.1	13	16	AAR88253
14	558	36.1	101	17	AAR95610
15	558	36.1	101	17	AAR95616
16	558	36.1	101	17	AAR97050
17	558	36.1	101	20	AAY44078
18	558	36.1	101	21	AAB18805
19	558	36.1	101	22	AAG67741
20	558	36.1	101	22	AAB84123
21	558	36.1	101	22	AAB68016
22	558	36.1	101	22	AAB67775
23	552.5	35.8	356	17	AAR95661
24	538	34.8	101	16	AAR88255
25	538	34.8	101	17	AAR95612
26	538	34.8	101	17	AAR95618
27	538	34.8	101	17	AAR97052
28	538	34.8	101	20	AAY44080
29	538	34.8	101	20	AAW97311
30	538	34.8	101	22	AAB84125
31	534	34.6	101	20	AAW97310
32	514	33.3	101	20	AAW97312
33	506	32.8	101	17	AAR95614
34	506	32.8	101	17	AAR97063
35	506	32.8	101	20	AAY44090
36	506	32.8	101	22	AAB84135
37	470.5	30.5	92	17	AAR95615
38	353	22.9	61	17	AAR97072
39	353	22.9	61	20	AAY44099
40	353	22.9	61	22	AAB84144
41	343	22.2	59	17	AAR97073
42	343	22.2	59	20	AAY44100
43	343	22.2	59	22	AAB84145
44	333	21.6	57	17	AAR97074
45	333	21.6	57	20	AAY44101

ALIGNMENTS

RESULT 1	
AAAR39286	
ID	AAAR39286 standard; Protein; 298 AA.
AC	AAAR39286;
DT	13-JAN-1994 (first entry)
DE	Respiratory syncytial virus (RSV) G protein.
KW	PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.
OS	Respiratory syncytial virus.
PN	W09314207-A.
XX	22-JUL-1993.
XX	05-JAN-1993; 93WC-CA00001.
XX	06-JAN-1992; 92GB-0000117.
XX	(CONN-) CONNAUGHT LAB LTD.
XX	Ewasysghyn ME, Klein MH;
XX	WPI, 1993-243222/30.
XX	N-PSDB: AAQ45686.
DR	Multimeric hybrid genes and their chimeric proteins - are
XX	PT vaccines against multiple pathogenic infections e.g.
PT	para-influenza virus and respiratory syncytial virus
XX	

PS Claim 11; Figure 7A-7D; 80pp; English.
 CC A novel multimeric hybrid gene is used as a vaccine. The gene
 CC consists of two gene sequences which are linked and encode antigenic
 CC regions, these two sequences being derived from two different
 CC pathogens (parainfluenza virus (PIV) and respiratory syncytial virus
 CC (RSV)). The gene sequences that are particularly used are those
 CC which encode PIV-3 F and HN proteins (AA045683, AA045684) and RSV F and
 CC G proteins (AA045685, AA045686).
 CC
 XX

Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 14; Length 298;
 Best Local Similarity 100.0%; Pred. No. 2.1e-113;

Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60
 DB 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60
 QY 61 FIASANHKTLLTALIODATSOIKNTPTPYLTODPOLGISFNSLSEITSOITTLASTP 120
 DB 61 FIASANHKTLLTALIODATSOIKNTPTPYLTODPOLGISFNSLSEITSOITTLASTP 120
 QY 121 GYKSNLOPTTVKTKNTTOTOPOPSKPTTKOROKPPKPNKPNDFEFVFNVPSCISNNP 180
 DB 121 GYKSNLOPTTVKTKNTTOTOPOPSKPTTKOROKPPKPNKPNDFEFVFNVPSCISNNP 180
 QY 181 TCWAICKRIIPNKKPKGKTTTPKPKPTTKKDLKPOTTKPREVPTTKPREPTINTK 240
 DB 181 TCWAICKRIIPNKKPKGKTTTPKPKPTTKKDLKPOTTKPREVPTTKPREPTINTK 240
 QY 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298
 DB 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 2

AAW96313
 ID AAW96313 standard; Protein; 298 AA.

AC AAW96313;

DT 28-JUN-1999 (first entry)

DE Membrane bound G protein of respiratory syncytial virus.

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

KW tissue plasminogen activator.

OS Respiratory syncytial virus.

PN WO9904010-A1.

PD 28-JAN-1999.

PF 16-JUL-1998; 98WO-CA00697.

PR 18-JUL-1997; 97US-0896442.

PA (CONN-) CONNAGHT LAB LTD.

PI Klein MH, Li X, Sambhara S;

DR WPI: 1999-132254/11.

DR N-PSDB; AAX08421.

XX Immunogenic composition for generating antibodies against
 PT respiratory syncytial virus - comprises non-replicating vector
 PT containing the protein G sequence, useful in protective vaccines and
 PT to raise antibodies for diagnosis
 XX

PS Claim 4; Fig 2; 67pp; English.

CC The respiratory syncytial virus (RSV) G protein can be used in
 CC vaccines by inserting the G protein gene into a non-replicating
 CC vector. The G protein is placed under the control of alternative
 CC signal and expression sequences, for example the chimeric G protein
 CC produced may also comprise the signal peptide of tissue plasminogen
 CC activator (tPA). The recombinant vector may also comprise sequences
 CC upstream of the G protein gene which enhance the G protein's
 CC immunoprotective ability. The resulting immunogenic composition will
 CC generate antibodies directed against the RSV G protein when
 CC administered to a host organism. The composition is useful as a
 CC vaccine to immunise against RSV-associated disease, particularly
 CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
 CC by usual immunisation and cell fusion methods.
 CC
 XX

Sequence 298 AA;

Query Match 100.0%; Score 1544; DB 20; Length 298;
 Best Local Similarity 100.0%; Pred. No. 2.1e-113;

Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60
 DB 1 MSKNKDORTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSILAMITSLITAI 60
 QY 61 FIASANHKTLLTALIODATSOIKNTPTPYLTODPOLGISFNSLSEITSOITTLASTP 120
 DB 61 FIASANHKTLLTALIODATSOIKNTPTPYLTODPOLGISFNSLSEITSOITTLASTP 120
 QY 121 GYKSNLOPTTVKTKNTTOTOPOPSKPTTKOROKPPKPNKPNDFEFVFNVPSCISNNP 180
 DB 121 GYKSNLOPTTVKTKNTTOTOPOPSKPTTKOROKPPKPNKPNDFEFVFNVPSCISNNP 180
 QY 181 TCWAICKRIIPNKKPKGKTTTPKPKPTTKKDLKPOTTKPREVPTTKPREPTINTK 240
 DB 181 TCWAICKRIIPNKKPKGKTTTPKPKPTTKKDLKPOTTKPREVPTTKPREPTINTK 240
 QY 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298
 DB 241 TTTTLLTNNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 3

AAP70845
 ID AAP70845 standard; protein; 298 AA.

AC AAP70845;

DT 05-APR-1991 (first entry)

DE Sequence of human respiratory syncytial virus (HRSV) A2 strain

DE G protein.

KW Vaccine.

OS Human respiratory syncytial virus (HRSV).

PN WO8704185-A.

PD 16-JUL-1987.

PF 23-DEC-1986; 86WO-US02756.

PR 14-JAN-1986; 86US-0818740.

PA (UNNC-) UNIV OF N CAROLINA.

DR (WERTZ) WERTZ G W.

DR (WERTZ) WERTZ G W.
 WPI: 1987-206300/29.
 DR N-PSDB; AAN70784.
 XX

PT Vaccines for human respiratory virus - comprising proteins or
 fragment encoded by a DNA sequence coding for human respiratory
 syncytial virus proteins.

XX Disclosure; Chart 13; 57pp; English.

CC A novel plasmid which comprises a DNA sequence encoding this
 CC protein, and the protein itself, are claimed, for use as HRSV
 CC vaccines. The vaccine can be administered to pregnant women or to
 CC women of child bearing age to stimulate maternal antibodies.
 CC Infants can also be vaccinated at 2-3 months of age.

XX Sequence 298 AA;

Query Match 91.8%; Score 1418; DB 8; Length 298;
 Best Local Similarity 93.3%; Pred. No. 1.6e-103;
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKKNKQRTAKTLEKTDNLHLLFISSGLYKLNKSVAGITLSILAMISTSLITTAII 60
 DB 1 MSKKNKQRTAKTLEKTDNLHLLFISSGLYKLNKSVAGITLSILAMISTSLITTAII 60
 QY 61 FIASANHKVTLTAIIQDATSOIKNTPTVLTQDPOLGISFNSLSEITTSOTTTILASTTP 120
 DB 61 FIASANHKVTLTAIIQDATSOIKNTPTVLTQDPOLGISFNSLSEITTSOTTTILASTTP 120
 QY 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPNNDFFHEFVNFVPCISCSNMP 180
 DB 121 GYKSTLOSTTVKTKNTTTQTOPSKPTTKORONKPPNNDFFHEFVNFVPCISCSNMP 180
 QY 181 TCWAICRIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTTKPTEEPPTINTTK 240
 DB 181 TCWAICRIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTTKPTEEPPTINTTK 240
 QY 241 TTTTTLNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298
 DB 241 TTTTTLNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

RESULT 4
 AAR25302

ID AAR25302 standard; Protein; 298 AA.

XX AAR25302;

DT 03-MAR-1993 (first entry)

DE HSRV glycoprotein G (gpc).

KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
 KW major capsid protein; N.

OS Human respiratory syncytial virus strain A2.

PN US5149650-A.

PD 22-SEP-1992.

PF 14-JAN-1986; 860S-0818740.

PR 14-JAN-1986; 860S-0818740.

PR 13-JUL-1988; 880S-0218737.

PA (UYNC-) UNIV NORTH CAROLINA.

PI Collins PL, Wertz GW;

DR WPI: 1992-340247/41.

PT N-PSDB; AA029623.
 PT Vaccines for human respiratory virus - include structural genes
 PT coding for native structural viral proteins and immunogenic
 PT fragments

XX Disclosure; Page 18; 21pp; English.

CC The sequences of mRNA encoding HRSV structural proteins are given in
 CC AA029622-26. The proteins are F, G, 22K, 9.5K and major capsid
 CC protein N. The sequences and encoded proteins are useful for
 CC preparing vaccines against HRSV. The vaccines can be used to confer
 CC immunity against respiratory tract infections on human subjects.

XX Sequence 298 AA;

Query Match 91.8%; Score 1418; DB 13; Length 298;
 Best Local Similarity 93.3%; Pred. No. 1.6e-103;
 Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKKNKQRTAKTLEKTDNLHLLFISSGLYKLNKSVAGITLSILAMISTSLITTAII 60
 DB 1 MSKKNKQRTAKTLEKTDNLHLLFISSGLYKLNKSVAGITLSILAMISTSLITTAII 60
 QY 61 FIASANHKVTLTAIIQDATSOIKNTPTVLTQDPOLGISFNSLSEITTSOTTTILASTTP 120
 DB 61 FIASANHKVTLTAIIQDATSOIKNTPTVLTQDPOLGISFNSLSEITTSOTTTILASTTP 120
 QY 121 GYKSNLOPTTVKTKNTTTQTOPSKPTTKORONKPPNNDFFHEFVNFVPCISCSNMP 180
 DB 121 GYKSTLOSTTVKTKNTTTQTOPSKPTTKORONKPPNNDFFHEFVNFVPCISCSNMP 180
 QY 181 TCWAICRIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTTKPTEEPPTINTTK 240
 DB 181 TCWAICRIPNKKPGKKTTPKPKPKPTTKKDLKDPOTTKREVPPTTKPTEEPPTINTTK 240
 QY 241 TTTTTLNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298
 DB 241 TTTTTLNNTGNPKLTSOMETFHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

RESULT 5
 AAM47605

ID AAM47605 standard; Protein; 298 AA.

XX AAM47605;

DT 11-JUN-1998 (first entry)

DE HRSV glycoprotein G.

KW HRSV; glycoprotein F; gpf; glycoprotein G; gpc; vaccine.

OS Human respiratory syncytial virus.

PN US5716823-A.

PD 10-FEB-1998.

PF 12-MAY-1997; 970S-0854783.

PR 13-JUL-1988; 880S-0218737.

PR 14-JAN-1986; 860S-0818740.

PR 23-DEC-1986; 86MO-US02756.

PR 11-JUN-1992; 920S-0897171.

PR 12-MAY-1997; 970S-0854783.

PA (PHAA) PHARMACIA & UPJOHN CO.

PI Collins PL, Wertz GW;

DR WPI: 1998-144802/13.
 DR N-PSDB; AAV18736.
 PT Production of human respiratory syncytial virus glyco-protein F or G
 PT - by culturing eukaryotic host cells transfected with corresponding
 PT DNA

PS Example 1; Columns 27-28; 17pp; English.

XX The present sequence was used in the development of a novel method

CC for the production of human respiratory syncytial virus (HRSV)

CC glycoprotein F (gpf) or glycoprotein G (gpc). The method comprises

CC culturing eukaryotic host cells transfected with an isolated DNA

CC sequence encoding HRSV gpf or gpc. The gp can be used to prepare

CC vaccines against HRSV.

XX

SQ Sequence 298 AA;

Query Match 91.8%; Score 1418; DB 19; Length 298;

Best Local Similarity 93.3%; Pred. No. 1.6e-103;

Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

DB 1 MSKNDORTAKTLEKMTDNLHLFISSGLYKLNKSAVOITLILAMISTSLIAI 60

QY 61 FIASANHKVTLTALIDATSOIKNTTPYLTQDPOLGISFSNISEITSOITTLASTP 120

DB 61 FIASANHKVTLTALIDATSOIKNTTPYLTQDPOLGISFSNISEITSOITTLASTP 120

QY 121 GYKSNLOPTTVTKNTTTOTOPSKPTTKORONKPPKPNDFEFVFNVPSCISNNP 180

DB 121 GYKSTLOSTTVTKNTTTOTOPSKPTTKORONKPPKPNDFEFVFNVPSCISNNP 180

QY 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

DB 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

QY 241 TITTTLLTNNTGNPKLTSMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

DB 241 TITTTLLTNNTGNPKLTSMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

RESULT 6

AAU74676 standard; Protein: 298 AA.

XX AAU74676;

AC

XX

DT 09-APR-2002 (first entry)

XX

DE Respiratory syncytial virus G protein.

XX

KW RSV; G protein; heavily glycosylated protein; antianaemic; antiviral;

KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;

KW antiviral chemotherapeutic compound; humoral response;

KW cellular immune response; HIV; paediatric respiratory disease;

KW globin gene transfer; sickle cell disease; beta-thalassaemia;

KW human immunodeficiency virus infection; HIV.

XX

OS Human respiratory syncytial virus.

XX

PN WO200192548-A2.

XX

PD 06-DEC-2001.

XX

PF 22-MAY-2001; 2001WO-US16610.

XX

PR 01-JUN-2000; 2000US-208701P.

XX

PA (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.

XX

PI Portner A, Takimoto T;

XX

DR MPI; 2002-130534/17.

XX

DR N-PSDB; AAS21045.

XX

PT Recombinant Sendai virus useful in vaccines to protect infection by

PT paramyxoviruses, comprises exogenous nucleic acid encoding

PT paramyxovirus protein or its antigenic fragment

XX

PS Disclosure; Page 48; 57pp; English.

XX

CC The invention relates to a recombinant Sendai virus comprising an

CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its

CC antigenic fragment. The virus may be administered in combination

CC with an antiviral chemotherapeutic compound. Two or more viruses

CC expressing different PMV proteins may be co-administered. Compositions

CC comprising the virus are useful for eliciting a humoral and/or

CC cellular immune response to a PMV in a mammal, particularly a human.

CC Further a recombinant Sendai virus comprising an exogenous nucleic acid

CC encoding a second PMV protein is also administered and priming and/or

CC boosting humoral or cellular immune response comprises administering

CC one or more of a recombinant or isolated PMV protein or its antigenic

CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral

CC vector encoding a PMV protein. The recombinant virus is useful as an

CC effective vaccine against HIV or RSV (the major causes of paediatric

CC respiratory disease) and also to express any gene of

CC interest in target cells, providing a positive medical impact on

CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)

CC into stem cells effects a cure for sickle cell disease or beta-

CC thalassaemia. The recombinant virus may also prove effective in

CC conferring immunity to human immunodeficiency virus (HIV) infection.

CC The Sendai virus replicates at level that is high enough to

CC induce sufficient immunity, but does not cause any harm to human

CC recipient. The present sequence represents a respiratory syncytial

CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein

CC suitable for expression by the recombinant virus of the invention.

XX

SQ Sequence 298 AA;

Query Match 91.8%; Score 1418; DB 23; Length 298;

Best Local Similarity 93.3%; Pred. No. 1.6e-103;

Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTLEKMTDNLHLFISSGLYKLNKSAVOITLILAMISTSLIAI 60

DB 1 MSKNDORTAKTLEKMTDNLHLFISSGLYKLNKSAVOITLILAMISTSLIAI 60

QY 61 FIASANHKVTLTALIDATSOIKNTTPYLTQDPOLGISFSNISEITSOITTLASTP 120

DB 61 FIASANHKVTLTALIDATSOIKNTTPYLTQDPOLGISFSNISEITSOITTLASTP 120

QY 121 GYKSNLOPTTVTKNTTTOTOPSKPTTKORONKPPKPNDFEFVFNVPSCISNNP 180

DB 121 GYKSTLOSTTVTKNTTTOTOPSKPTTKORONKPPKPNDFEFVFNVPSCISNNP 180

QY 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

DB 181 TCWAICKRIIPNKKPGKKTTPKPTKPKTKKDLKPOTTKPEVPTTKPEEPTINTTK 240

QY 241 TITTTLLTNNTGNPKLTSMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

DB 241 TITTTLLTNNTGNPKLTSMETPHSTSEGNLSPSOVSTSEHPSPSPPTTRQ 298

RESULT 7

AAW96314 standard; Protein: 232 AA.

XX AAW96314;

AC

XX

DT 28-JUN-1999 (first entry)

XX

DE Secreted G protein of respiratory syncytial virus.

XX

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

KW tissue plasminogen activator.

XX

OS Respiratory syncytial virus.

XX

PN WO9904010-A1.

```

XX 28-JAN-1999.
PD 16-JUL-1998; 98WO-CA00697.
XX 18-JUL-1997; 97US-0896442.
PR (CONN-) CONNAUGHT LAB LTD.
XX Klein MH, Li X, Sambhara S;
XX WPI: 1999-132254/11.
DR N-PSDB: AAX08422.
XX
PT Immunogenic composition for generating antibodies against
PT respiratory syncytial virus - comprises non-replicating vector
PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis
XX
PS Claim 9; Fig 3; 67pp; English.
XX
CC The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternative
CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods. This truncated G
CC protein is secreted since it lacks a transmembrane domain.
XX
SQ Sequence 232 AA:
XX
Query Match 79.6%; Score 1229; DB 20; Length 232;
Best Local Similarity 100.0%; Pred. No. 8 3e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 67 HKVTLTTAIIQDATSOIKNTPTVLTDPDLGISFSNLSSEITSTQTTILASTTPGVKSNL 126
DB 1 HKVTLTTAIIQDATSOIKNTPTVLTDPDLGISFSNLSSEITSTQTTILASTTPGVKSNL 60
QY 127 OPTVTKTKNTTTOTOPSPKPTTKORONKPPNNDPHEFVNFVPCSCSNPTCWAIC 186
DB 61 OPTVTKTKNTTTOTOPSPKPTTKORONKPPNNDPHEFVNFVPCSCSNPTCWAIC 120
QY 187 KRIIPNKKPGKKTPTTKPTKPTTKKDLKPTQTKPKKEVPTTKPEEPTINTTKNTITTT 246
DB 121 KRIIPNKKPGKKTPTTKPTKPTTKKDLKPTQTKPKKEVPTTKPEEPTINTTKNTITTT 180
QY 247 LITNNTGNPKLTSQMEFTHSTSSSEGNLSPSOVSTTSEHPSSSPPTTRQ 298
DB 181 LITNNTGNPKLTSQMEFTHSTSSSEGNLSPSOVSTTSEHPSSSPPTTRQ 222

RESULT 8
AAP90441
ID AAP90441 standard; protein: 681 AA.
XX
AC AAP90441;
XX
DT 01-NOV-1989 (first entry)
XX
DE Chimeric human respiratory syncytial virus glycoproteins F and G.
XX
KW Chimeric polypeptide: human respiratory syncytial virus;
XX protein F; protein G; vaccine.
OS Human respiratory syncytial virus.

```

```

XX MO8905823-A.
PN 29-JUN-1989.
XX
XX 31-OCT-1988; 88WO-US03784.
XX 23-DEC-1987; 87US-0137387.
XX
XX (UPJO ) UPJOHN CO.
XX
XX Wathen M;
XX
DR WPI: 1989-206593/28.
XX
PT Chimeric human respiratory syncytial virus polypeptides(s)
PT - contg. immunogenic fragments from HNSV glycoproteins
PT F and G, for vaccine prodn.
XX
PS Claim 3; page 47-48; 50pp; English.
XX
CC Chimeric polypeptide contg. a signal sequence and one or more
CC immunogenic fragments from both human respiratory syncytial virus
CC glycoproteins F and G. Can be used in vaccines. Hosts are, eg
CC E. coli, Chinese hamster ovary cells, murine C127 cells and
CC S. frugiperda.
XX
SQ Sequence 681 AA:
XX
Query Match 58.4%; Score 901; DB 10; Length 681;
Best Local Similarity 91.0%; Pred. No. 1.7e-62;
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
QY 96 QLGISFSNLSRITQOTTTILASTTPGVKSNLOPTVTKTKNTTTOTOPSPKPTTKORONK 155
DB 490 QLGISFSNLSRITQOTTTILASTTPGVKSNLOPTVTKTKNTTTOTOPSPKPTTKORONK 549
QY 156 PNKPNNDFHEFVNFVPCSCSNPTCWAICRIPNKKPGKKTPTTKPTKPTTKKDL 215
DB 550 PSKPNNDPHEFVNFVPCSCSNPTCWAICRIPNKKPGKKTPTTKPTKPTTKKDP 609
QY 216 KPQTTKPKKEVPTTKPEEPTINTTKNTITTTLLTNNTGNPKLTSQMEFTHSTSSSEGNLS 275
DB 610 KPQTTKPKKEVPTTKPEEPTINTTKNTITTTLLTNNTGNPKLTSQMEFTHSTSSSEGNLS 669
QY 276 PSQVSTTSE 284
DB 670 PSQVNTISSQ 678

RESULT 9
AAB68336
ID AAB68336 standard; protein: 299 AA.
XX
AC AAB68336;
XX
DT 09-JUL-2001 (first entry)
XX
DE Amino acid sequence of RSV G-protein.
XX
KW Respiratory syncytial virus; RSV; G-protein; annexin II; L-selectin;
XX RSV infection.
XX
OS Respiratory syncytial virus.
XX
PN WO200129054-A2.
XX
PD 26-APR-2001.
XX
PE 23-OCT-2000; 2000WO-GB04084.
XX
PR 21-OCT-1999; 99GB-0024990.
XX

```


PT esp. by conjugation with new *Klebsiella pneumoniae* p40 protein
XX
PS Disclosure: Page 23; 38pp; French.
XX
CC The sequence of the wild type amino acid residues 130-230 from the
CC protein G of the respiratory syncytial virus (RSV) subgroups A. The
CC sequence was used to synthesize the immunogenic peptides AAR8245-52
CC based on residues 174-187 of the subgroup A and B RSVs. The peptides are
CC pref. conjugated to a novel carrier protein (the p40 protein; see
CC AAR8237) derived from a membrane lipopolysaccharide (LPS) fraction from
CC *Klebsiella pneumoniae*. The LPS fraction was isolated by precipitating the
CC *Klebsiella* membrane LPSs with a divalent cation and detergents,
CC subjecting the recovered proteins to anion-exchange chromatography to
CC obtain an immunological adjuvant and linking the p40 protein to the above
CC peptides. The conjugates are useful in the treatment of RSV A or B
CC infections.
XX
SQ Sequence 101 AA;
Query Match 36.1%; Score 558; DB 16; Length 101;
Best Local Similarity 99.0%; Pred. No. 1.5e-36;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 130 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVNFVPCISCSNPTCAICKRI 189
DB 1 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVNFVPCISCSNPTCAICKRI 60
OY 190 PNKKPGKTTTKPKKPTFKTKDKLPQTTRKREVPPTTKP 230
DB 61 PNKKPGKTTTKPKKPTFKTKDKHQPQTTRKREVPPTTKP 101
RESULT 14
AAR95610
ID AAR95610 standard; peptide; 101 AA.
XX
AC AAR95610;
XX
DT 11-FEB-1997 (first entry)
XX
DE RSV subgp. A protein G wild type amino acids 130-230.
XX
KW Respiratory syncytial virus; protein G; heterologous protein; infection;
KW cell surface; *Staphylococcus xylosum*; *Staphylococcus carnosus*; vaccine;
KW expression vector; shuttle vector; hydrophobic region; immunogenic;
XX streptococcal protein G serum albumin binding domain; conformation.
XX
OS Synthetic.
XX
PN WO9614409-A1.
XX
PD 17-MAY-1996.
XX
PF 07-NOV-1995; 95WO-FR01464.
XX
PR 07-NOV-1994; 94FR-0013307.
XX
PA (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI Binz H, Nguyen Ngoc T, Nygren PA, Stahl S, Uhlen M;
XX
XX WPI; 1996-251759/25.
DR N-PSDB; AAT27066.
XX
PT Secretion of recombinant hydrophobic peptide analogue - esp. for
PT secretion of respiratory syncytial virus antigenic peptide(s)
PT suitable for use in oral vaccines
XX
PS Claim 12; Page 20; 47pp; French.
XX
CC This is the sequence of the wild type protein G amino acid 130-230 from
CC the respiratory syncytial virus (RSV) subgroup A. The sequence is used
CC in a method for producing heterologous proteins on the surface of a cell

CC pref. a non-pathogenic micro-organism such as *Staphylococcus xylosum* or
CC *S. carnosus*. The coding sequence was generated synthetically by
CC annealing and ligating the oligonucleotides AAT4252-74. The complete
CC ligated sequence was inserted into the expression shuttle vector pSE/BRXM
CC such that the G2 sequence (encoding amino acids 130-230) is fused to the
CC sequence encoding the streptococcal protein G serum albumin binding
CC domain (BB). The encoded peptides may have modifications in their
CC hydrophobic regions which do not affect their activity but allow them to
CC traverse the cell membrane and be displayed on the cell surface in the
CC correct immunogenic conformations. Cells carrying such peptides can be
CC used as vaccines against RSV infections.
XX
SQ Sequence 101 AA;
Query Match 36.1%; Score 558; DB 17; Length 101;
Best Local Similarity 99.0%; Pred. No. 1.5e-36;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 130 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVNFVPCISCSNPTCAICKRI 189
DB 1 TVTKNTTTTQTOPSKPTTKORONKPPNNDHFEEVNFVPCISCSNPTCAICKRI 60
OY 190 PNKKPGKTTTKPKKPTFKTKDKLPQTTRKREVPPTTKP 230
DB 61 PNKKPGKTTTKPKKPTFKTKDKHQPQTTRKREVPPTTKP 101
RESULT 15
AAR95616
ID AAR95616 standard; Protein; 101 AA.
XX
AC AAR95616;
XX
DT 10-FEB-1997 (first entry)
XX
DE RSV sub-group A wild type protein G residues 130-230.
XX
KW Heterologous protein; cell surface; *Staphylococcus xylosum*; protein G;
KW *Staphylococcus carnosus*; respiratory syncytial virus; wild type; mutant;
KW expression plasmid; fusion protein; streptococcal.
XX
OS Human respiratory syncytial virus.
XX
PN WO9614418-A1.
XX
PD 17-MAY-1996.
XX
PF 07-NOV-1995; 95WO-FR01465.
XX
PR 07-NOV-1994; 94FR-0013309.
XX
PA (FABR) FABRE MEDICAMENT SA PIERRE.
XX
PI Binz H, Nguyen Ngoc T, Nygren PA, Stahl S, Uhlen M;
XX
XX WPI; 1996-251768/25.
DR N-PSDB; AAT27073.
XX
PT Expression of respiratory syncytial virus protein G at cell surface
PT - of bacteria non-pathogenic for mammals, useful in orally active
PT vaccines
XX
PS Claim 12; Page 13-14; 38pp; French.
XX
CC The invention relates to the prodn. of a heterologous protein on the
CC surface of a cell, pref. *Staphylococcus xylosum* or *S. carnosus*, contg. a
CC fragment of the respiratory syncytial virus comprising residues 130-230
CC of the RSV protein G. This sequence is through to be the wild type
CC residues 130-230 from the RSV subgp. A. The heterologous protein may be
CC also contain the subgp. B sequence (AAT27074). These sequences may be
CC mutated to replace the Cys residues at pos. 173 and 186 by residues
CC unable to form a disulphide bridge esp. Ser residues (see AAT27075-6).
CC Other mutations include substitution of the Phe residues at pos. 163,

CC 165, 168 and/or 170 by Ser residues. The sequences encoding the wild
 CC type or mutant RSV residues 130-230 are inserted into plasmid
 CC PSE'mpl8BBXM to create plasmid PSE/G2BBXM (wild type sequence) or
 CC PSE/G2subBBXM (mutant sequence). The sequence is produced from PSE/BBXM
 CC as a fusion protein where the fusion is with the streptococcal protein G
 CC serum albumin binding domain. These plasmids are introduced into
 CC S. xyloosus or S. carnosus for expression of the wild type or mutant
 CC protein.
 CC
 XX

SO Sequence 101 AA:

Query Match 36.1%; Score 558; DB 17; Length 101;
 Best Local Similarity 99.0%; Pred. No. 1.5e-36;
 Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 130 TVTKNTTTTQTOPSKPTTKORONKPPNKNDFHFVNFVPCSTCSNPTCWAICKRI 189
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 TVTKNTTTTQTOPSKPTTKORONKPPNKNDFHFVNFVPCSTCSNPTCWAICKRI 60
 OY 190 PNKKPGKTTTTPPKKPTFKTKKDLKPOTTKPREVPTTKP 230
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 PNKKPGKTTTTPPKKPTFKTKKDLKPOTTKPREVPTTKP 101

Search completed: May 1, 2003, 13:44:22
 Job time : 38.6717 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:43:28 ; Search time 15.7434 Seconds
(Without alignments)
556.934 Million cell updates/sec

Title: US-09-462-816-2
Perfect score: 1544
Sequence: 1 MSKNKDQRTAKTEKTDL.....VSTSPHPSQSPSPVTRQ 298

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued_Patents_AA:*
2: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/BACKFILE1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1544	100.0	298	2	US-08-467-963C-8
2	1544	100.0	298	2	US-08-838-189D-8
3	1544	100.0	298	3	US-08-852-344D-8
4	1544	100.0	298	3	US-08-344-639E-8
5	1544	100.0	298	4	US-08-467-969A-8
6	1544	100.0	298	4	US-08-467-961A-8
7	1544	100.0	298	4	US-08-001-554A-8
8	901	58.4	681	6	5194595-19
9	293.5	19.1	263	5	PCT-US91-08177-13
10	222	14.4	37	3	US-08-793-792-12
11	193	12.5	32	3	US-08-793-792-8
12	188	12.2	216	3	US-08-928-361B-8
13	188	12.2	1837	3	US-08-928-361B-5
14	187.5	12.1	1721	3	US-08-700-651-5
15	187.5	12.1	1721	3	US-08-928-361B-6
16	183.5	11.9	216	3	US-08-928-361B-27
17	174	11.3	28	3	US-08-793-792-4
18	173.5	11.2	249	3	US-08-700-651-15
19	173.5	11.2	249	3	US-08-928-361A-20
20	169	10.9	36	4	US-09-082-279B-871
21	169	10.9	36	4	US-09-315-304B-871
22	164.5	10.7	2476	2	US-08-276-967-2
23	162	10.5	941	4	US-07-757-022B-14
24	162	10.5	1022	4	US-07-757-022B-84
25	162	10.5	1038	4	US-07-757-022B-74
26	162	10.5	1049	4	US-07-757-022B-58
27	162	10.5	1140	4	US-07-757-022B-104

28	162	10.5	1270	4	US-07-757-022B-44	Sequence 44, Appl
29	162	10.5	1311	4	US-07-757-022B-42	Sequence 42, Appl
30	162	10.5	1313	4	US-07-757-022B-142	Sequence 142, App
31	162	10.5	1314	4	US-07-757-022B-50	Sequence 50, Appl
32	162	10.5	1320	4	US-07-757-022B-46	Sequence 46, Appl
33	162	10.5	1320	4	US-07-757-022B-60	Sequence 60, Appl
34	162	10.5	1354	4	US-07-757-022B-48	Sequence 48, Appl
35	162	10.5	1361	4	US-07-757-022B-40	Sequence 40, Appl
36	162	10.5	1363	4	US-07-757-022B-52	Sequence 52, Appl
37	162	10.5	1404	4	US-07-757-022B-2	Sequence 2, Appl
38	162	10.5	1404	4	US-07-757-022B-62	Sequence 62, Appl
39	161.5	10.5	878	4	US-09-556-706B-2	Sequence 2, Appl
40	159.5	10.3	907	3	US-08-783-774-2	Sequence 2, Appl
41	159.5	10.3	907	4	US-09-328-599A-1	Sequence 1, Appl
42	159.5	10.3	907	5	PCT-US95-04611A-19	Sequence 19, Appl
43	158	10.2	175	3	US-08-700-651-12	Sequence 12, Appl
44	158	10.2	175	3	US-08-928-361B-17	Sequence 17, Appl
45	151	9.8	37	3	US-08-793-792-11	Sequence 11, Appl

ALIGNMENTS

RESULT 1
US-08-467-963C-8
; Sequence 8, Application US/08467963C
; Patent No. 5968776
GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; TITLE OF INVENTION: EMASYSN, Mary E
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Slim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,963C
; FILING DATE:
CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/838,189
; FILING DATE: 16-APR-1997
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
; NAME: STEWART, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-467-963C-8

Query Match 100.0%; Score 1544; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTALIIDATSOIKNTPTVLTODPOLGISFSNLSSETTSQTTTILASTP 120
DB 61 FIASANHKVLTALIIDATSOIKNTPTVLTODPOLGISFSNLSSETTSQTTTILASTP 120
QY 121 GVSNNLOPTVTKNTTTOTOPSKPTTKORONKPPKPNNDPFHEVNFVPCISCSNPP 180
DB 121 GVSNNLOPTVTKNTTTOTOPSKPTTKORONKPPKPNNDPFHEVNFVPCISCSNPP 180
QY 181 TCMAICRIPNKKPGKKTTPKPKPTKTKDLKPKQTTPKREVPPTTKPEEPTINTTK 240
DB 181 TCMAICRIPNKKPGKKTTPKPKPTKTKDLKPKQTTPKREVPPTTKPEEPTINTTK 240
QY 241 TTTTTLTNNNTGNPKLTISOMETFHSTSEGNLSPSOVSTTSEHPSPSSPPTTRQ 298
DB 241 TTTTTLTNNNTGNPKLTISOMETFHSTSEGNLSPSOVSTTSEHPSPSSPPTTRQ 298

RESULT 2

US-08-838-189D-8
Sequence 8, Application US/08838189D
Patent No. 5998169
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-APR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-838-189D-8

Query Match 100.0%; Score 1544; DB 2; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
DB 1 MSKNDQRTAKTLEKTDWTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTALIIDATSOIKNTPTVLTODPOLGISFSNLSSETTSQTTTILASTP 120
DB 61 FIASANHKVLTALIIDATSOIKNTPTVLTODPOLGISFSNLSSETTSQTTTILASTP 120
QY 121 GVSNNLOPTVTKNTTTOTOPSKPTTKORONKPPKPNNDPFHEVNFVPCISCSNPP 180
DB 121 GVSNNLOPTVTKNTTTOTOPSKPTTKORONKPPKPNNDPFHEVNFVPCISCSNPP 180
QY 181 TCMAICRIPNKKPGKKTTPKPKPTKTKDLKPKQTTPKREVPPTTKPEEPTINTTK 240
DB 181 TCMAICRIPNKKPGKKTTPKPKPTKTKDLKPKQTTPKREVPPTTKPEEPTINTTK 240
QY 241 TTTTTLTNNNTGNPKLTISOMETFHSTSEGNLSPSOVSTTSEHPSPSSPPTTRQ 298
DB 241 TTTTTLTNNNTGNPKLTISOMETFHSTSEGNLSPSOVSTTSEHPSPSSPPTTRQ 298

RESULT 3

US-08-852-344D-8
Sequence 8, Application US/08852344D
Patent No. 6017539
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
TITLE OF INVENTION: SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,344D
FILING DATE: 07-MAY-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/344,639
FILING DATE: 14-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single

```
; TOPOLOGY: linear
; US-08-852-344D-8
Query Match          100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKQRTAKTLEKTWDTLNLHLLFISSGLYKLNLSVAQITLSILAMISTSLITAI 60
Db 1 MSKNKQRTAKTLEKTWDTLNLHLLFISSGLYKLNLSVAQITLSILAMISTSLITAI 60
Qy 61 FIASANHKVLTITAIIDATSIQIKNTPTTYLTQDPGLGISFSLSEITSCQTTTILASTP 120
Db 61 FIASANHKVLTITAIIDATSIQIKNTPTTYLTQDPGLGISFSLSEITSCQTTTILASTP 120
Qy 121 GVSNLQPTVTKTKNTTTTQTOPSKPTTKOROKNPKPNNDPHEFVFNVPKCSICSNP 180
Db 121 GVSNLQPTVTKTKNTTTTQTOPSKPTTKOROKNPKPNNDPHEFVFNVPKCSICSNP 180
Qy 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKKDLKPQTTKPKKEVPTTKPTEEPITNTTK 240
Db 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKKDLKPQTTKPKKEVPTTKPTEEPITNTTK 240
Qy 241 TINTTTLLTNTGNTGNPKLTSQMETFHSISSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298
Db 241 TINTTTLLTNTGNTGNPKLTSQMETFHSISSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298

RESULT 4
US-08-344-639E-8
; Sequence 8, Application US/08344639E
; Patent No. 6033668
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,639E
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-391 MIS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 8:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-344-639E-8
Query Match          100.0%; Score 1544; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSKNKQRTAKTLEKTWDTLNLHLLFISSGLYKLNLSVAQITLSILAMISTSLITAI 60
Db 1 MSKNKQRTAKTLEKTWDTLNLHLLFISSGLYKLNLSVAQITLSILAMISTSLITAI 60
Qy 61 FIASANHKVLTITAIIDATSIQIKNTPTTYLTQDPGLGISFSLSEITSCQTTTILASTP 120
Db 61 FIASANHKVLTITAIIDATSIQIKNTPTTYLTQDPGLGISFSLSEITSCQTTTILASTP 120
Qy 121 GVSNLQPTVTKTKNTTTTQTOPSKPTTKOROKNPKPNNDPHEFVFNVPKCSICSNP 180
Db 121 GVSNLQPTVTKTKNTTTTQTOPSKPTTKOROKNPKPNNDPHEFVFNVPKCSICSNP 180
Qy 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKKDLKPQTTKPKKEVPTTKPTEEPITNTTK 240
Db 181 TCWAICKRIPKKRGKKTTPPKPKPTFKTKKDLKPQTTKPKKEVPTTKPTEEPITNTTK 240
Qy 241 TINTTTLLTNTGNTGNPKLTSQMETFHSISSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298
Db 241 TINTTTLLTNTGNTGNPKLTSQMETFHSISSEGNLSPSOVSTTSEHPSOSPSPNTTRQ 298

RESULT 5
US-08-467-969A-8
; Sequence 8, Application US/08467969A
; Patent No. 6168786
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,969A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:dh
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
```

TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-969A-8

Query Match 100.0%; Score 1544; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTEKTDNLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
|||||
DB 1 MSKNDQRTAKTEKTDNLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTITAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTILLASTP 120
|||||
DB 61 FIASANHKVLTITAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTILLASTP 120
QY 121 GVSANLOPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNPP 180
|||||
DB 121 GVSANLOPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNPP 180
QY 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKKEVPTTKPEEPTINTTK 240
|||||
DB 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKKEVPTTKPEEPTINTTK 240
QY 241 TTTTTLTNTTGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298
|||||
DB 241 TTTTTLTNTTGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298

RESULT 6

US-08-467-961A-8
Sequence 8, Application US/08467961A
Patent No. 6171783

GENERAL INFORMATION:

APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasashyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6TH FLOOR

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,961A

FILING DATE: 06-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/001,554

FILING DATE: 06-JAN-1993

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9200117.1

FILING DATE: 06-JAN-1992

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-476 Mib:hb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-961A-8

Query Match 100.0%; Score 1544; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTEKTDNLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
|||||
DB 1 MSKNDQRTAKTEKTDNLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTITAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTILLASTP 120
|||||
DB 61 FIASANHKVLTITAIIDATSOIKNTPTTYLTODPOLGISFSNLSEITSTILLASTP 120
QY 121 GVSANLOPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNPP 180
|||||
DB 121 GVSANLOPTTVKTKNTTTTQTOPSKPTTKORONKPPNNDFHFEVNFVPCISCSNPP 180
QY 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKKEVPTTKPEEPTINTTK 240
|||||
DB 181 TCMAICRIPNKKPGKKTTPKTKPKTKKDKLPQTKRKKEVPTTKPEEPTINTTK 240
QY 241 TTTTTLTNTTGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298
|||||
DB 241 TTTTTLTNTTGNPKLTSQMETFHSTSSSEGNLSPSQVSTTSEHPSQSSPPTTRQ 298

RESULT 7

US-08-001-554A-8

Sequence 8, Application US/08001554A
Patent No. 6225091

GENERAL INFORMATION:

APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasashyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/001,554A

FILING DATE: 06-JAN-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-286

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

TELEX: 065-24567 SIMBAS

```
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE:
US-08-001-554A-8
```

```
Query Match          100.0%; Score 1544; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1,7e-132;
Matches 298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MSKNKDORTAKTLEKTDLTINHLFISSGLYKLNKSAQITLSIAMIISTLITAI 60
Db 1 MSKNKDORTAKTLEKTDLTINHLFISSGLYKLNKSAQITLSIAMIISTLITAI 60
QY 61 FIASANHKVLTITAIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTP 120
Db 61 FIASANHKVLTITAIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTP 120
QY 121 GKSNIQPTTVKTKNTTQTQPSKPTTKORONKPPNKNNDFEFVNFVPCISNNP 180
Db 121 GKSNIQPTTVKTKNTTQTQPSKPTTKORONKPPNKNNDFEFVNFVPCISNNP 180
QY 181 TCWAICRIIPNKKPGKKTTPKPKPTEKTKKDLKPQTKPEKVPPTKPTPEPTINTTK 240
Db 181 TCWAICRIIPNKKPGKKTTPKPKPTEKTKKDLKPQTKPEKVPPTKPTPEPTINTTK 240
QY 241 TTTTTLTNNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSSPNTTRQ 298
Db 241 TTTTTLTNNNTGNPKLTSOMETHSTSSSEGNLSPSOVSTTSEHPSSPNTTRQ 298
```

```
RESULT 8
5194595-19
Patent No. 5194595
Applicant: WATHEM, MICHAEL W.
TITLE OF INVENTION: CHIMERIC GLYCOPROTEINS CONTAINING
IMMUNOGENIC SEGMENT OF THE GLYCOPROTEINS OF HUMAN RESPIRATORY
SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/543,780
FILING DATE: 31-OCT-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 137,387
FILING DATE: 23-DEC-1987
SEQ ID NO: 19
LENGTH: 681
5194595-19
```

```
Query Match          58.4%; Score 901; DB 6; Length 681;
Best Local Similarity 91.0%; Pred. No. 1.3e-73;
Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;
```

```
QY 96 QLGISFNSLSEITSTIILASTIPGVKSNLOPTVTKNTTQTQPSKPTTKORONK 155
Db 490 QLGISFNSLSEITSTIILASTIPGVKSNLOPTVTKNTTQTQPSKPTTKORONK 155
QY 156 PNKPNNDHFVFNFPVPCISNNPQWAIKRIIPNKKPGKKTTPKPKPTTKKDL 215
Db 550 PSKRNDHFVFNFPVPCISNNPQWAIKRIIPNKKPGKKTTPKPKPTTKKDL 215
QY 216 KPQTTKPEVPTKPTPEPTINTTKNTITTLTNNNTGNKRLTSOMETHSTSSSEGNL 275
Db 610 KPQTTKPEVPTKPTPEPTINTTKNTITTLTNNNTGNKRLTSOMETHSTSSSEGNL 275
QY 276 PSQVSTSE 284
Db 670 PSQVSTSE 284
```

```
RESULT 9
PCT-US91-08177-13
Sequence 13, Application PC/TUS9108177
GENERAL INFORMATION:
Applicant: Samal, Siba K
TITLE OF INVENTION: Bovine Respiratory Syncytial Virus Genes
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Venable, Baetjer, Howard & Civiletti
STREET: 1201 New York Avenue N.W., suite 1000
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08177
FILING DATE: 19911104
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/608,937
FILING DATE: 05-NOV-1990
ATTORNEY/AGENT INFORMATION:
NAME: Higne, David W
REGISTRATION NUMBER: 30,265
REFERENCE/DOCKET NUMBER: 20509-96711
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-962-4854
TELEFAX: 202-962-8300
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 263 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US91-08177-13
```

```
Query Match          19.1%; Score 295.5; DB 5; Length 263;
Best Local Similarity 31.7%; Pred. No. 3.2e-19;
Matches 86; Conservative 35; Mismatches 123; Indels 27; Gaps 7;
```

```
QY 1 MSKNKDORTAKTLEKTDLTINHLFISSGLYKLNKSAQITLSIAMIISTLITAI 60
Db 1 MSKNKDORTAKTLEKTDLTINHLFISSGLYKLNKSAQITLSIAMIISTLITAI 60
QY 61 FIASANHKVLTITAIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTP 120
Db 61 FIASANHKVLTITAIQDATSQIKNTPTTYLQDPOLGISFNSLSEITSTIILASTP 120
QY 121 GVK-----SNLOPTVTKNTT-TTQTOPSKPTTKORONKPPNKPNNDHFVNFVPC 174
Db 119 GTTYGHPINTQNNKIKISQSTPLATKRLPINP-----ESNPNENHDHNNSTLPHVPC 174
QY 175 TCSNPNPCWAIK-----KRIPN-----KKPKKTTTPKPKPTTKKDLKPQTKP 222
Db 175 TCSNPNPCWAIK-----KRIPN-----KKPKKTTTPKPKPTTKKDLKPQTKP 222
QY 223 KEVPTTKPTPEPTINTTKNTITTLTNNNT 253
Db 235 TATPQGGILSSP---EHQTNOSTQISQHTS 262
```

```
RESULT 10
US-08-793-792-12
Sequence 12, Application US/08793792
Patent No. 6077511
GENERAL INFORMATION:
Applicant:
TITLE OF INVENTION: Antigenic peptides derived from the
```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESS: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: VERNY, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1837 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-928-361B-5

Query Match 12.2%; Score 188; DB 3; Length 1837;
Best Local Similarity 22.7%; Pred. No. 2.8e-08;
Matches 70; Conservative 32; Mismatches 161; Indels 46; Gaps 5;

QY 9 TAKTLEKTDWTLNHLLEISSGLYKLNLSVAQITLSIAMIIST----- 52
DB 239 TGTIPNTY-----AGYRSNENKTEPESANTNPLVDPRKNPCSENSFEVQ 288
QY 53 -----SLITTAIFIASNHKVTLLTAIIQATSOIKNTPTPYLLQDPOLGISFSNLSI 107
DB 289 IFDMGSKYIPIYTKCVGKHTTTTTTTTTTTTTTTTTTTTTTTT-----TTTTT 339
QY 108 TSQTLLASTTPGVKSNLQPTVTKNTTTTQOPSKPTTKOKONKPNKPNDFHEV 167
DB 340 TT- 397
QY 168 FNFVPCISCSNPNPCMAICKRIPNKKPKTKPTTKPKTKKDLKPOTKPREVPT 227
DB 398 TT 457
QY 228 TKPPEEPIINTKNTITTTLLTNNTGNPKLTSOMETFSSTSSSEGNLSPSOVSTSEHP 287
DB 458 TTATTTTKKPTT---TTTTTTT 513
QY 288 QPSSPNTT 296
DB 514 TTTTATTTT 522

RESULT 14
US-08-700-651-5
Sequence 5, Application US/08700651B
Patent No. 6015882
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.

APPLICANT: GUT, JIRI
TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAs AND RNAs
TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
TITLE OF INVENTION: INFECTIONS
FILE REFERENCE: 480.19-4(HV)
CURRENT APPLICATION NUMBER: US/08/700,651B
CURRENT FILING DATE: 1997-08-14
EARLIER APPLICATION NUMBER: 08/415,751
EARLIER FILING DATE: 1995-04-03
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 1721
TYPE: PRT
ORGANISM: Cryptosporidium parvum
US-08-700-651-5

Query Match 12.1%; Score 187.5; DB 3; Length 1721;
Best Local Similarity 25.6%; Pred. No. 2.8e-08;
Matches 61; Conservative 21; Mismatches 125; Indels 31; Gaps 2;

QY 67 HKVTLTAIIQDATSQIKNTPTPYLLQDPOLGISFSNLSITQTTLASTTPGVKSNL 126
DB 191 HTT 250
QY 127 QPTVTKNTTTTQTOPSKPTTKORONKPNKPNDFEVNFVPCISCSNPNPCMAIC 186
DB 251 TTTTTTTTTTTTTTTTTTTTTTTT-----TTTTTTTTTTTTTT 287
QY 187 KRIPNKKRGKTTTKPKTKPTTKKDLKPOTKPREVPT-----KPEEPIINT 238
DB 288 TT 347
QY 239 TKNTITTTLLTNNTGNPKLTSOMETFSSTSSSEGNLSPSOVSTSEHPQSSPNTT 296
DB 348 TTATTTT 405

RESULT 15
US-08-928-361B-6
Sequence 6, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESS: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: VERNY, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.76-1(HV)
TELECOMMUNICATION INFORMATION:

TELEPHONE: 650-324-1677
 TELEFAX: 650-324-1678
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1721 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-928-361B-6

Query Match 12.1%; Score 187.5; DB 3; Length 1721;
 Best Local Similarity 25.6%; Pred. No. 2.8e-08;
 Matches 61; Conservative 21; Mismatches 125; Indels 31; Gaps 2;

QY 67 HKVLTALIDATSOIKNTPTVLTODPQIGISFNSLSEITSOITTLASTTPEGVKSNL 126
 | | | | | : | | | | : | | | | :
 DB 191 HHTT 250

QY 127 QPTVKTNTTTOPOSKPTTKORONKPNKPNNDPHEVENFVPCISCSNPTCWAIC 186
 | | | | | : | | | | : | | | | :
 DB 251 TT 287

QY 187 KRIPNKKPGKTTTKPKPFKTKKDLKQDTKPKVEPTT-----KPIEPTINT 238
 | | | | | : | | | | : | | | | :
 DB 288 TT 347

QY 239 TKTNTITTLTNNITGNPKLTQSOMETFHSSTSEGNLSPQVSTSEHPSQSPSPNTT 296
 | | | | | : | | | | : | | | | :
 DB 348 TT 405

Search completed: May 1, 2003, 13:48:02
 Job time : 17.2434 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:46:33 ; Search time 77.5925 Seconds

(Without alignments)
331.391 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDQRTAKTEKTWDTL.....VSTTSEHPSSPPTTRQ 298

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA:*

1: /cgn2-6/ptodata/2/pubppaa/US08_NEW_PUB pep:.*
2: /cgn2-6/ptodata/2/pubppaa/PCT_NEW_PUB pep:.*
3: /cgn2-6/ptodata/2/pubppaa/US06_NEW_PUB pep:.*
4: /cgn2-6/ptodata/2/pubppaa/US06_PUBCOMB pep:.*
5: /cgn2-6/ptodata/2/pubppaa/US07_NEW_PUB pep:.*
6: /cgn2-6/ptodata/2/pubppaa/US07_PUBCOMB pep:.*
7: /cgn2-6/ptodata/2/pubppaa/PCTUS_PUBCOMB pep:.*
8: /cgn2-6/ptodata/2/pubppaa/US08_PUBCOMB pep:.*
9: /cgn2-6/ptodata/2/pubppaa/US09_NEW_PUB pep:.*
10: /cgn2-6/ptodata/2/pubppaa/US09_PUBCOMB pep:.*
11: /cgn2-6/ptodata/2/pubppaa/US10_NEW_PUB pep:.*
12: /cgn2-6/ptodata/2/pubppaa/US10_PUBCOMB pep:.*
13: /cgn2-6/ptodata/2/pubppaa/US60_NEW_PUB pep:.*
14: /cgn2-6/ptodata/2/pubppaa/US60_PUBCOMB pep:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	193	12.5	5179	9	US-10-025-380-1068
2	193	12.5	5179	10	US-09-922-217-1068
3	193	12.5	5179	10	US-09-833-263-1068
4	169.5	11.0	1367	10	US-09-801-368-108
5	162.5	10.5	288	10	US-09-216-393-341
6	162.5	10.5	288	10	US-09-216-393-344
7	162	10.5	941	12	US-10-124-557-14
8	162	10.5	1022	12	US-10-124-557-84
9	162	10.5	1038	12	US-10-124-557-74
10	162	10.5	1049	12	US-10-124-557-58
11	162	10.5	1140	12	US-10-124-557-104
12	162	10.5	1270	12	US-10-124-557-44
13	162	10.5	1311	12	US-10-124-557-42
14	162	10.5	1313	12	US-10-124-557-142
15	162	10.5	1314	12	US-10-124-557-50
16	162	10.5	1320	12	US-10-124-557-46
17	162	10.5	1320	12	US-10-124-557-60
18	162	10.5	1354	12	US-10-124-557-48
19	162	10.5	1361	12	US-10-124-557-40

20	162	10.5	1363	12	US-10-124-557-52	Sequence 52, Appl
21	162	10.5	1404	12	US-10-124-557-2	Sequence 2, Appl
22	162	10.5	1404	12	US-10-124-557-62	Sequence 62, Appl
23	150.5	9.7	995	9	US-09-984-130-48	Sequence 48, Appl
24	145	9.4	1075	10	US-09-801-368-110	Sequence 110, Appl
25	144.5	9.4	1322	10	US-10-176-847-54	Sequence 114, Appl
26	144.5	9.4	2828	10	US-09-905-129-21	Sequence 54, Appl
27	144.5	9.4	2828	10	US-09-991-630-21	Sequence 21, Appl
28	144.5	9.4	2828	10	US-10-028-072-224	Sequence 224, Appl
29	144	9.3	449	9	US-10-121-049-224	Sequence 224, Appl
30	144	9.3	449	9	US-10-123-904-224	Sequence 224, Appl
31	144	9.3	449	9	US-10-140-474-224	Sequence 224, Appl
32	144	9.3	449	9	US-10-140-474-224	Sequence 224, Appl
33	144	9.3	449	9	US-10-175-746-224	Sequence 224, Appl
34	144	9.3	449	9	US-10-176-921-224	Sequence 224, Appl
35	144	9.3	449	9	US-10-176-921-224	Sequence 224, Appl
36	144	9.3	449	9	US-10-137-865-224	Sequence 224, Appl
37	144	9.3	449	9	US-10-140-474-224	Sequence 224, Appl
38	144	9.3	449	9	US-10-142-431-224	Sequence 224, Appl
39	144	9.3	449	9	US-10-143-114-224	Sequence 224, Appl
40	144	9.3	449	9	US-10-140-002-224	Sequence 224, Appl
41	144	9.3	449	9	US-10-142-419-224	Sequence 224, Appl
42	144	9.3	449	9	US-10-123-262-224	Sequence 224, Appl
43	144	9.3	449	9	US-10-142-423-224	Sequence 224, Appl
44	144	9.3	449	9	US-10-121-050-224	Sequence 224, Appl
45	144	9.3	449	9	US-10-141-755-224	Sequence 224, Appl

ALIGNMENTS

RESULT 1
US-10-025-380-1068
Sequence 1068, Application US/10025380
Publication NO. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Lodes, Michael J.
APPLICANT: Secrist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Meagher, Madeleine Joy
APPLICANT: Stolk, John A.
APPLICANT: Wang, Yunglong
APPLICANT: Jiang, Yunglong
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasir A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick Thomas S.
APPLICANT: Carter, Darrick
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025,380
CURRENT FILING DATE: 2001-12-19
NUMBER OF SEQ ID NOS: 1129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1068
LENGTH: 5179
TYPE: PRT
ORGANISM: Homo sapiens
US-10-025-380-1068

Query Match 12.5%; Score 193; DB 9; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.7e-06;

Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 71 LTVATIDDA--TSQIKWTP-----TYLQDPLGTSFNSLSEIT--SQTTLATTPQ 121
DB 1504 WTTPTTPASTTTLPTTSSPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563

```
QY 122 ----VKSNIOPITVKTKNTTT-----OTQPSKPTTKORONKPKPNNDH 164
      :      :      :      :      :      :      :      :      :
Db 1564 PPPTTTTTPPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1623
      :      :      :      :      :      :      :      :      :
QY 165 FEVFNFPQSGICSNPPCAICRIKPKKGGKTTTKPKTKPKTKDKLQOTKPKKE 224
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TTPSTITLTP-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
      :      :      :      :      :      :      :      :      :
QY 225 VPTTKPEEPTITNTKNTITNTLTNNTPGNPKLTOSMETHSSEGNLSPQSVSTSE 284
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTTPMTTTPSPPTTTP--SSPITTTTTPSSSTTTPSPPTTTPMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 285 HPSQPSPPNTT 296
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTTPSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 2
US-09-922-217-1068
; Sequence 1068, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-922-217-1068

Query Match 12.5%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.7e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;
```

```
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeline J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-833-263-1068

Query Match 12.5%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.7e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 71 LTTAIIODA--TSQIKNTTP-----TYLQDPOLGISFNSLSEIT--SQTTTILASTTPG 121
      :      :      :      :      :      :      :      :      :
Db 1504 MTTPITPPASTTTLPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563
      :      :      :      :      :      :      :      :      :
QY 122 ----VKSNIOPITVKTKNTTT-----OTQPSKPTTKORONKPKPNNDH 164
      :      :      :      :      :      :      :      :      :
Db 1564 PPPTTTPPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1623
      :      :      :      :      :      :      :      :      :
QY 165 FEVFNFPQSGICSNPPCAICRIKPKKGGKTTTKPKTKPKTKDKLQOTKPKKE 224
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TTPSTITLTP-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
      :      :      :      :      :      :      :      :      :
QY 225 VPTTKPEEPTITNTKNTITNTLTNNTPGNPKLTOSMETHSSEGNLSPQSVSTSE 284
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTTPMTTTPSPPTTTP--SSPITTTTTPSSSTTTPSPPTTTPMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 285 HPSQPSPPNTT 296
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTTPSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 4
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 1367
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

Query Match
Best Local Similarity 11.0%; Score 169.5; DB 10; Length 1367;
Best Local Similarity 26.2%; Pred. No. 8.5e-05;
Matches 62; Conservative 36; Mismatches 94; Indels 45; Gaps 8;

QY 84 KNTTPYVLQDPOLGISFNLSITSTOTLLASTPGV--KSNLOPTTVKKNIT--T 139
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 209 KSTSTTSSTSESTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 268
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 140 QTPSKPTTKORONKPPNPNNDHFVFNPCSCSNPPCMAICKRIPNKKGKKT 199
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 269 KEKPTPTTSCKEKPTPPHD-----TPPCT-----KKTTTSKTCCKT 310
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 200 TKPTKPTKTKKDLK-----QTKPEVPTTKPTEE-----PINTTKTNTTTL 247
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 311 TTPVPPTSSSTESSAPVPTPSSSTESSAPVTSSSTESSAPVPTPSSSTESSAP 370
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 248 LNNNTGNGKILNSQMETFHSTSEGNLSPQVSTSEHSP-----SSPPNT 296
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 371 VVSSIT---ESSAPVTSSTESSAPVPTPSSSTESSAPVTSSSTESSAPVTS 424
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

RESULT 5
US-09-216-393-341
; Sequence 341 Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 341
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393-341

Query Match
Best Local Similarity 10.5%; Score 162.5; DB 10; Length 288;
Best Local Similarity 23.1%; Pred. No. 4.5e-05;
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;

QY 39 AQTLSITAMISTLITAIIFIASANKVLTALIIDATSOI----- 83
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 76 AALSDAIDRVSDPLVSLDIVIREAAQAKFDLGRLLTIDIASIGEGAMALMGEAAFI 135
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 84 -----KNTTPYVLQDPOLGISFNLSITSTOTLLASTPGVKSNILOPTTVKTKN 135
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 136 RPRRSKRGGKKT---TSSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 181
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 136 TTTTQTSKPTTKORONKPPNPNNDHFVFNPCSCSNPPCMAICKRIPNKKRG 195
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 182 TTTTSTTTTPTT----- 194
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 196 KTTTKPKPTKPTTKKDLKPKQTTPKPEVPTTKPEEPTINTTKTNTTLLTNNNTGN 255
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 195 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTP 249
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 256 PKLTSQMETFHSTSEGNLSPQVSTSEHSPSSPPNT 296
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 250 PTTT-----TTTEPTTTTSTTTTTTTTTTTTTTPTSTTTST 284
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
```

```
RESULT 6
US-09-216-393-344
; Sequence 344 Application US/09216393
; Patent No. US20010014447A1
; GENERAL INFORMATION:
; APPLICANT: Milhausen, Michael James
; TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
; FILE REFERENCE: TX-1-C2
; CURRENT APPLICATION NUMBER: US/09/216,393
; CURRENT FILING DATE: 1998-12-18
; EARLIER APPLICATION NUMBER: 08/994,825
; EARLIER FILING DATE: 1997-12-19
; NUMBER OF SEQ ID NOS: 364
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 344
; LENGTH: 288
; TYPE: PRT
; ORGANISM: Toxoplasma gondii
US-09-216-393-344

Query Match
Best Local Similarity 10.5%; Score 162.5; DB 10; Length 288;
Best Local Similarity 23.1%; Pred. No. 4.5e-05;
Matches 65; Conservative 22; Mismatches 99; Indels 95; Gaps 5;

QY 39 AQTLSITAMISTLITAIIFIASANKVLTALIIDATSOI----- 83
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 76 AALSDAIDRVSDPLVSLDIVIREAAQAKFDLGRLLTIDIASIGEGAMALMGEAAFI 135
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 84 -----KNTTPYVLQDPOLGISFNLSITSTOTLLASTPGVKSNILOPTTVKTKN 135
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 136 RPRRSKRGGKKT---TSSSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTSTST 181
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 136 TTTTQTSKPTTKORONKPPNPNNDHFVFNPCSCSNPPCMAICKRIPNKKRG 195
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 182 TTTTSTTTTPTT----- 194
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 196 KTTTKPKPTKPTTKKDLKPKQTTPKPEVPTTKPEEPTINTTKTNTTLLTNNNTGN 255
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 195 -TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTP 249
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 256 PKLTSQMETFHSTSEGNLSPQVSTSEHSPSSPPNT 296
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 250 PTTT-----TTTEPTTTTSTTTTTTTTTTTTTTPTSTTTST 284
| : : | : : | : : | : : | : : | : : | : : | : : | : : |
| : : | : : | : : | : : | : : | : : | : : | : : | : : |

RESULT 7
US-10-124-557-14
; Sequence 14 Application US/10124557
; Patent No. US20020137894A1
; GENERAL INFORMATION:
; APPLICANT: Turner, Katherine
; Clark, Stephen C.
; Jacobs, Kenneth
; Hewick, Rodney M.
; Gesner, Thomas G.
; TITLE OF INVENTION: Megakaryocyte Stimulating Factors
; NUMBER OF SPOUNCES: 143
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/124,557
```



```
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ. ID NO: 74:
SEQUENCE CHARACTERISTICS:
LENGTH: 1038 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 74:
US-10-124-557-74

Query Match 10.5%; Score 162; DB 12; Length 1038;
Best Local Similarity 24.6%; Pred. No. 0.00023;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDATSOIKNTPTTYITODPOLGISFSNLSITSTTTTILASTTGGVKSNILOPTTV 131
DB 190 TSKVLAKPTPKAEYTK-----GPAITPKPEPTPTTKPEPASVTP--KEPTPTTI 237
QY 132 K-----TKNTTTTOTOPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSN 178
DB 238 KSAPTTKPEAPPTTKSAPTTKPEAPPTTKPEAPTTKPEP-----APTITKEP 286
QY 179 NPTGMAICKRIPN-----KKPGKTTTKP-----TKKPTFKTTK 212
DB 287 APTTKSAPTTKPEAPPTTKPKAPPTTKPEAPTTKPEPTTKPEAPTTKPEAPTTKPK 346
QY 213 KDL-----KPTTKPKPEVPTTKP-----TEEPTINTKINTTTLLNNNTGNPKLIS 260
DB 347 EPAPTPAKPKAPPTTKPEAPTTKPEAPTTTKESPTTKPEAPPTTKSAPTTTKPEAPT 406
QY 261 QMETFHSSTSEGNLSPSOVSTTSEHPSOSPSPNTT 296
DB 407 TTKSAPTTKPE-----PSPTTITKPEAPTTTKPEAPTT 438

RESULT 10
US-10-124-557-58
: Sequence 58, Application US/10124557
: Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
```

```
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSER: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Cserr, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ. ID NO: 58:
SEQUENCE CHARACTERISTICS:
LENGTH: 1049 amino acids
TYPE: amino acid
MOLECULE TYPE: protein
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-10-124-557-58

Query Match 10.5%; Score 162; DB 12; Length 1049;
Best Local Similarity 24.6%; Pred. No. 0.00023;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

QY 72 TTAIIQDATSOIKNTPTTYITODPOLGISFSNLSITSTTTTILASTTGGVKSNILOPTTV 131
DB 233 TSKVLAKPTPKAEYTK-----GPAITPKPEPTPTTKPEPASVTP--KEPTPTTI 280
QY 132 K-----TKNTTTTOTOPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSN 178
DB 281 KSAPTTKPEAPPTTKSAPTTKPEAPTTTKPEAPTTKPEP-----APTITKEP 329
QY 179 NPTGMAICKRIPN-----KKPGKTTTKP-----TKKPTFKTTK 212
DB 330 APTTKSAPTTKPEAPPTTKPKAPPTTKPEAPTTKPEPTTKPEAPTTKPEAPTTKPK 389
QY 213 KDL-----KPTTKPKPEVPTTKP-----TEEPTINTKINTTTLLNNNTGNPKLIS 260
DB 390 EPAPTPAKPKAPPTTKPEAPTTKPEAPTTTKESPTTKPEAPPTTKSAPTTTKPEAPT 449
QY 261 QMETFHSSTSEGNLSPSOVSTTSEHPSOSPSPNTT 296
DB 450 TTKSAPTTKPE-----PSPTTITKPEAPTTTKPEAPTT 481

RESULT 11
US-10-124-557-104
: Sequence 104, Application US/10124557
: Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
```


Thu May 1 14:50:33 2003

us-09-462-816-2.rapb

Page 8

```

QY 179 NPYCMALCKRIEIN-----KKPKKITTTP-----TKRPTEPTTK 212
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 330 APPTTKSAPPTPKKBPAPPTPKKBPAPPTPKKBPAPPTPKKBPAPTTTK 389
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 213 KDL-----KPTTKRKEVPTTKP-----TEEPTINTTKINTITLLNNTGNPKLIS 260
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 390 EPAPAPAPKBPAPTTPKKBPAPTTPKKBPAPTTTKESPTTKKBPAPTTK 449
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 261 QMETPHSTSSSEGNLSPQSVSTTSEHPQSPSPPTT 296
      :: : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 450 TTKSAPTTPKP-----PSPPTTKKBPAPTTPKKBPAPTT 481
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15
US-10-124-557-50
; Sequence 50, Application US/10124557
; Patent No. US20020137894A1

GENERAL INFORMATION:
APPLICANT: Turner, Katherine

APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewlick, Rodney M.
Gesner, Thomas G.

TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive

CITY: Cambridge

STATE: Massachusetts
COUNTRY: U.S.A.

ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.00

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/124,557

FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502

FILING DATE: 18-JAN-1991

APPLICATION NUMBER: US 07/546,114

FILING DATE: 29-JUN-1990

APPLICATION NUMBER: US 01/451,196

FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901

FILING DATE: 08-AUG-1985

ATTORNEY/AGENT INFORMATION:

NAME: CSELF, LUANN
DESTINATION NUMBER: 31 833

REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GT 5190

REFERENCE/DOC# NUMBER:

TELEPHONE: (617) 876-1170

TELEPHONE: (617) 876-5881
TELEFAX: (617) 876-5882

INFORMATION FOR SEO ID NO: 50:

EXPERIENCE CHARACTERISTICS:

LENGTH: 1314 amino acids

TYPE: amino acid

TOPOLOGY: linear

ECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 50:

4-557-50

Query Match	Score	DB	Length
10.58;	162;	12;	1314;

Best Local Similarity 24.6%; Pred. No. 0.00037

Matches	68;	Conservative	26;	Mismatches	104;	Indels	10;	Gaps	2
---------	-----	--------------	-----	------------	------	--------	-----	------	---

72 TTAIQDATSQIKNTPTYLQDPQGISFSNLSEIISQITWLLASITPGVKSNLQPIIV 13

```

Db 234 TSKVLAPRTKAEITTK-----GPAITTEKEPTPTTKBPASATP---KEPPTTI 281
QY 132 K-----TKNTTTTQOPSPAPTTKOKONKRPKNPNNDPHEVNEFAPCSGCSN 178
Db 282 KSAETTPKEBPAPTTTKSAPTTKBPAPTTKEBPAPTTKBPED-----APTTTKE 330
QY 179 NPTGMAICAKRIPN-----KKPKKTKTTRP-----TKPKTEKTK 212
Db 331 APITTKSAPTPKEBPAPTTTPKPKAPTTTPKEBPAPTTKEBPAPTTKEBPAPTTK 350
QY 213 KDL-----KQPTTKPKVEVPTTK-----TEBPTINTTKNTLTTLTNTGNPRLTS 260
Db 351 EPAPTPAPKRPAPTTKBPAPTTTPKEBPAPTTTKESPTTPKBPAPTTTKSAPTTTKEBAP 450
QY 261 OMETFHSTFSGEGLNSQVSTTSEHPQSOPSSPNTT 296
Db 451 TTKSAPTTPKR-----PSPTTKKEBPAPTTTPKBPAPTT 482

```

Search completed: May 1, 2003, 13:56:34
Job time : 82.5925 secs

JOB TIME : 02.3920 SECS

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:41:33 ; Search time 19.6792 Seconds
(without alignments)
1455.750 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544
Sequence: 1 MSKNKDDRTAKTLKMTDL.....VSTSEHPSQSPPTTRQ 298

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1530	99.1	298	1 MGNZRL	major surface glycoprotein
2	1418	91.8	298	1 MGNZ	major surface glycoprotein
3	1343	87.0	297	2 J01205	attachment protein
4	1343	87.0	297	2 J01208	attachment protein
5	1333	86.3	298	2 J05680	G protein - Human
6	1328	86.0	297	2 J01209	attachment protein
7	1324	85.8	297	2 J01206	attachment protein
8	1313	85.0	298	2 J01207	attachment protein
9	1307	84.7	297	2 J01204	attachment protein
10	767.5	49.7	292	1 MGNZ60	major surface glycoprotein
11	761	49.3	292	1 MGNZ18	major surface glycoprotein
12	425	27.5	467	1 VHNZ	nucleocapsid protein
13	299	19.4	250	2 P00768	glycoprotein G - b
14	299	19.4	257	1 MGNZBR	glycoprotein G - b
15	295.5	19.1	263	2 J02284	glycoprotein G - b
16	284	18.1	263	2 J02388	glycoprotein G - b
17	280	18.1	248	2 P00769	glycoprotein G - b
18	272	17.6	263	2 A48732	attachment glycoprotein
19	208	13.5	307	1 GSFF3	salivary glue protein
20	194	12.6	662	2 A45155	mucin FIM-C.1 - Af
21	193	12.5	3020	2 A43932	mucin 2 precursor, hypothetical
22	191	12.4	851	2 T22696	hypothetical protein
23	191	12.4	1161	2 S57180	probable membrane
24	190	12.3	322	2 A53715	apomucin precursor
25	188	12.2	1832	2 T31113	mucin-like glycoprotein
26	185.5	12.0	379	2 S50125	larval glue protein
27	183	11.9	217	2 S01358	salivary glue protein
28	183	11.9	354	2 T46740	microfilament sheath
29	177	11.5	770	2 T22808	hypothetical protein

30	175	11.3	825	2 T29634	hypothetical protein
31	172.5	11.2	327	2 S20074	promastigote surface
32	172.5	11.2	1777	2 T34369	hypothetical protein
33	171.5	11.1	216	2 T51920	mucin - rhesus mac
34	170.5	11.0	279	2 S53363	mucin 5AC (clone J
35	170	11.0	3570	2 T45025	mucin MUC5B, trache
36	169.5	11.0	1367	1 S48478	glucan 1,4-alpha-g
37	168	10.9	402	2 E86185	hypothetical prote
38	167.5	10.8	294	2 A37232	mucin, tracheal (A
39	167.5	10.8	1630	2 A53577	ascites staloglyco
40	164.5	10.7	2476	2 T34022	zonadhesin - pig
41	163	10.6	371	2 S20075	promastigote surfa
42	163	10.6	477	2 S53362	mucin 5AC (clone J
43	160.5	10.4	393	2 S62335	171-7 protein - fr
44	160.5	10.4	866	2 T45462	membrane glycoprot
45	160.5	10.4	867	2 T43463	membrane glycoprot

ALIGNMENTS

RESULT 1
MGNZRL
major surface glycoprotein G - human respiratory syncytial virus (strain Long)
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: A32703, S12279
R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.
Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987
A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and
A:Reference number: A32703; MUID:87289657; PMID:2441388
A:Accession: A32703
A:Molecule type: mRNA
A:Residues: 1-298 <DOR>
A:Cross-References: GB:M17212; NID:9333940; PIDN:AAA47411.1; PID:9333941
R:Garcia-Barreno, B.; Portella, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.
EMBO J. 9, 4181-4187, 1990
A:Title: Frame shift mutations as a novel mechanism for the generation of neutralizat
A:Reference number: S12279; MUID:91065351; PMID:2249671
A:Accession: S12279
A:Molecule type: mRNA
A:Residues: 1-298 <GAR>
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:41-63/Domain: transmembrane #status predicted <TMN>
F:85,103,135,179,237,250,251,273,294/Binding site: carbohydrate (Asn) (covalent) #sta

Query Match 99.1%; Score 1530; DB 1; Length 298
Best local similarity 99.0%; Pred. No. 2.4e-92;
Matches 295; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MSKNKDDRTAKTLKMTDLNLHLFISSGLYKLKLSVAQITTSILAMISTSLITAIT 60
DB 1 MSKNKDDRTAKTLKMTDLNLHLFISSGLYKLKLSVAQITTSILAMISTSLITAIT 60

QY 61 FIAANHKVLTITTAIIDATSOIKNTPTVLTDDPOLGISFSNSETTSOTTTILASTP 120
DB 61 FIAANHKVLTITTAIIDATSOIKNTPTVLTDDPOLGISFSNSETTSOTTTILASTP 120

QY 121 GVKSNLOPTVTKTKNTTTTQTPSKPTTKORONKPPKPNNDHFVEYFNVPSCISNNP 180
DB 121 GVKSNLOPTVTKTKNTTTTQTPSKPTTKORONKPPKPNNDHFVEYFNVPSCISNNP 180

QY 121 GVKSNLOPTVTKTKNTTTTQTPSKPTTKORONKPPKPNNDHFVEYFNVPSCISNNP 180
DB 121 GVKSNLOPTVTKTKNTTTTQTPSKPTTKORONKPPKPNNDHFVEYFNVPSCISNNP 180

QY 181 TCMAICRIPKPKGKTK 240
DB 181 TCMAICRIPKPKGKTK 240

QY 241 TMTTTLTNNNTGNPKLTSGMETHSTSEGNLSPSOVSTSEHPSQSPPTTRQ 298
DB 241 TMTTTLTNNNTGNPKLTSGMETHSTSEGNLSPSOVSTSEHPSQSPPTTRQ 298

RESULT 2

MG22
Major surface glycoprotein G - human respiratory syncytial virus
C:Species: human respiratory syncytial virus
C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 24-Sep-1999
C:Accession: A94048; A93599; A04039
R:Wertz, G.W.; Collins, P.L.; Huang, Y.; Gruber, C.; Levine, S.; Ball, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 4075-4079, 1985
A:Title: Nucleotide sequence of the G protein gene of human respiratory syncytial virus
A:Reference number: A94048; M01D:85216636; PMID:3858665
A:Accession: A94048
A:Molecule type: mRNA
A:Residues: 1-298 <MR>
A:Cross-references: GB:K01486; GB:K02719; GB:K03348; GB:K03349; GB:M1217; GH
7.1; PID:9333932
A:Note: residues 207-298 are identical with residues 376-467 of the nucleocapsid protein
A:Note: this protein may carry 40-80 separate O-linked carbohydrate chains distributed a
R:Satake, M.; Colligan, J.E.; Elango, N.; Morley, E.; Venkatesan, S.
Nucleic Acids Res. 13, 7795-7812, 1985
A:Title: Respiratory syncytial virus envelope glycoprotein (G) has a novel structure.
A:Reference number: A93599; M01D:86067198; PMID:4069997
A:Accession: A93599
A:Molecule type: mRNA
A:Residues: 1-298 <SAT>
A:Cross-references: GB:X03149; M1D:960997; PIDN:CA26928.1; PID:960998
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:38-66/Domain: transmembrane #status predicted <TNM>
F:85,135,237,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 91.8%; Score 1418; DB 1; Length 298;
Best Local Similarity 93.3%; Pred. No. 4,4e-85;
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITSLIAMIISTSLITAI 60
DB 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITSLIAMIISTSLITAI 60
OY 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSOITTLASTP 120
DB 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSOITTLASTP 120
OY 121 GVSANLOPTVTKNTTQTOTOPSKPTTKORONKPPNPNNDHFVFNVPSCISNPP 180
DB 121 GVSANLOPTVTKNTTQTOTOPSKPTTKORONKPPNPNNDHFVFNVPSCISNPP 180
OY 181 TCWAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTKPKKEVPTTKPEEPTINTTK 240
DB 181 TCWAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTKPKKEVPTTKPEEPTINTTK 240
OY 241 TMTITLLTNTGPKLTISOMETPHSTSEGNLSPSOVSTTSEHPSSPSPMTTR 298
DB 241 TMTITLLTNTGPKLTISOMETPHSTSEGNLSPSOVSTTSEHPSSPSPMTTR 298

RESULT 3
J01205
Attachment protein - human respiratory syncytial virus (strain RSB1734)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: J01205
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A
A:Reference number: J01204; M01D:91374005; PMID:1895054
A:Accession: J01205
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill
children and adults.
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 87.0%; Score 1343; DB 2; Length 297;
Best Local Similarity 88.6%; Pred. No. 3.2e-80;
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

OY 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITSLIAMIISTSLITAI 60
DB 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITSLIAMIISTSLITAI 60
OY 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSOITTLASTP 120
DB 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSOITTLASTP 120
OY 121 GVSANLOPTVTKNTTQTOTOPSKPTTKORONKPPNPNNDHFVFNVPSCISNPP 180
DB 121 GVSANLOPTVTKNTTQTOTOPSKPTTKORONKPPNPNNDHFVFNVPSCISNPP 180
OY 181 TCWAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTKPKKEVPTTKPEEPTINTTK 240
DB 181 TCWAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTKPKKEVPTTKPEEPTINTTK 240
OY 241 TMTITLLTNTGPKLTISOMETPHSTSEGNLSPSOVSTTSEHPSSPSPMTTR 296
DB 241 TMTITLLTNTGPKLTISOMETPHSTSEGNLSPSOVSTTSEHPSSPSPMTTR 296

RESULT 4
J01208
Attachment protein - human respiratory syncytial virus (strain RSB6256)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: J01208
R:Cane, P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:Title: Identification of variable domains of the attachment (G) protein of subgroup
A:Reference number: J01204; M01D:91374005; PMID:1895054
A:Accession: J01208
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract
children and adults.
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 87.0%; Score 1343; DB 2; Length 297;
Best Local Similarity 88.5%; Pred. No. 3.2e-80;
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

OY 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITSLIAMIISTSLITAI 60
DB 1 MSKNKQDRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITSLIAMIISTSLITAI 60
OY 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSOITTLASTP 120
DB 61 FIASANKHVTLTATLIDATSOIKNTPTTYLTQDPOLGISFNSLSEITSOITTLASTP 120
OY 121 GVSANLOPTVTKNTTQTOTOPSKPTTKORONKPPNPNNDHFVFNVPSCISNPP 180
DB 121 GVSANLOPTVTKNTTQTOTOPSKPTTKORONKPPNPNNDHFVFNVPSCISNPP 180
OY 181 TCWAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTKPKKEVPTTKPEEPTINTTK 240
DB 181 TCWAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTKPKKEVPTTKPEEPTINTTK 240
OY 241 TMTITLLTNTGPKLTISOMETPHSTSEGNLSPSOVSTTSEHPSSPSPMTTR 296
DB 241 TMTITLLTNTGPKLTISOMETPHSTSEGNLSPSOVSTTSEHPSSPSPMTTR 296

RESULT 5
J05680

G protein - human respiratory syncytial virus
C:Species: Human respiratory syncytial virus
C:Date: 11-Nov-1997 #sequence_revision 11-Nov-1997 #text_change 26-Feb-1998
C:Accession: J05680
R:Gen: X.; Wang, Z.; Qian, Y.; Zhu, R.; Deng, J.; Du, J.; Zhu, Z.
Chinese J. Virol. 12, 317-322, 1996
A:title: Molecular analysis of G protein gene of a respiratory syncytial virus strain 15
A:Reference number: J05680
A:Accession: J05680
A:Molecule type: mRNA
A:Residues: 1-298 <GEN>
A:Experimental source: strain B79
A:Note: the authors translated the codon TTT for residue 165 and 170 as Glu, TTC for res
F:1-38/Domain: Intracellular #status predicted <INT>
F:39-66/Domain: Intracellular #status predicted <INT>
F:67-298/Domain: Extracellular #status predicted <EXC>

Query Match 86.3%; Score 1333; DB 2; Length 298;
Best Local Similarity 87.9%; Pred. No. 1.4e-79;
Matches 262; Conservative 10; Mismatches 26; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
1 MSKTKDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
Db 1 MSKTKDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
QY 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
QY 121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
Db 121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
QY 181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
Db 181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
QY 241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 298
241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 298
Db 241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 298

RESULT 6

J01209
attachment protein - human respiratory syncytial virus (strain RSB614)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: J01209
R:Can: P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:title: Identification of variable domains of the attachment (G) protein of subgroup A
A:Reference number: J01204; MUID:91374005; PMID:1895054
A:Accession: J01209
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status predict

Query Match 86.0%; Score 1328; DB 2; Length 297;
Best Local Similarity 87.8%; Pred. No. 3e-79;
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
1 MSKTKDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
Db 1 MSKTKDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
QY 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120

Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
QY 121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
Db 121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
QY 181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
Db 181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
QY 241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 296
241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 296
Db 241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 296

RESULT 7

J01206
attachment protein - human respiratory syncytial virus (strain RSB5857)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: J01206
R:Can: P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:title: Identification of variable domains of the attachment (G) protein of subgroup
A:Reference number: J01204; MUID:91374005; PMID:1895054
A:Accession: J01206
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,251,273,294/Binding site: carbohydrate (Asn) (covalent) #status pred

Query Match 85.8%; Score 1324; DB 2; Length 297;
Best Local Similarity 87.8%; Pred. No. 5.5e-79;
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
1 MSKTKDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
Db 1 MSKTKDORTAKTEKTDNLHLFISSGLYKLNKSAQITLSILAMITSLITTAII 60
QY 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
Db 61 FIASANKVLTTLAIIDATSOIKNTPTTYLTQDPOLGIFSNSLSEITSTTTILASTP 120
QY 121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
Db 121 GKSNLOPTTVKTKNTTQOPSKPTTKORONKPPNNDFHEVFNVPSCISNNP 180
QY 181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
Db 181 TCMAICRIIPNKKPGKKTTPKPKPTTKKDLKPQTTPKKEVPTTKPTEPTINTTK 240
QY 241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 296
241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 296
Db 241 TTTTLLTNNTGNPKLTSQMETFHSTSSGNSLSPSOVSTTSEHPSPSPPTNTTQ 296

RESULT 8

J01207
attachment protein - human respiratory syncytial virus (strain RSB6190)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: J01207
R:Can: P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:title: Identification of variable domains of the attachment (G) protein of subgroup
A:Reference number: J01204; MUID:91374005; PMID:1895054
A:Accession: J01207
A:Molecule type: mRNA

A:Residues: 1-298 <CAN>
A:Note: the authors translated the codon ACC for residue 4 as Asn and AGC for residue 22
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill-
idren and adults.
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,237,250,273,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 85.0%; Score 1313; DB 2; Length 298;
Best Local Similarity 86.9%; Pred. No. 2,8e-78;
Matches 259; Conservative 11; Mismatches 28; Indels 0; Gaps 0;

Db 1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
1 MSKTDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
Oy 61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
Db 61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
Oy 121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
Db 121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
Oy 181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 240
181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 240
Db 181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 240
Oy 241 TMTITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298
241 TMTITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298
Db 241 TMTITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 9
J01204
attachment protein - human respiratory syncytial virus (strain NS642)
N:Alternate names: G protein
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997
C:Accession: J01204
R:Canine, P.A.; Matthews, D.A.; Pringle, C.R.
J. Gen. Virol. 72, 2091-2096, 1991
A:Title: Identification of variable domains of the attachment (G) protein of subgroup A
A:Reference number: J01204; MUID:91374005; PMID:18959054
A:Accession: J01204
A:Molecule type: mRNA
A:Residues: 1-297 <CAN>
C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill-
idren and adults.
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:85,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 84.7%; Score 1307; DB 2; Length 297;
Best Local Similarity 86.5%; Pred. No. 6,9e-78;
Matches 257; Conservative 12; Mismatches 28; Indels 0; Gaps 0;

Db 1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
Oy 1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
Db 1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
Oy 61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
Db 61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
Oy 121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
Db 121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
Oy 181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 240
181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 240
Db 181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 240

Oy 241 TMTITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 297
241 TMTITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 297
Db 241 TMTITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 297

RESULT 10
MG218
major surface glycoprotein G - human respiratory syncytial virus (strain 8/60)
N:Alternate names: attachment glycoprotein G
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: A37077
R:Sullender, W.M.; Anderson, K.; Wertz, G.W.
Virology 178, 195-203, 1990
A:Title: The respiratory syncytial virus subgroup B attachment glycoprotein: analysis
gous subgroup virus challenge
A:Reference number: A37077; MUID:90357765; PMID:1697126
A:Accession: A37077
A:Molecule type: mRNA
A:Residues: 1-292 <SUL>
A:Cross-references: EMBL:M55633; NID:g333944; PIDN:AAA47413.1; PID:g333945
C:Genetics:
A:Gene: G
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:45-63/Domain: transmembrane #status predicted <TMN>
F:81,86,100,220,290/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.7%; Score 767.5; DB 1; Length 292;
Best Local Similarity 55.3%; Pred. No. 6,5e-43;
Matches 162; Conservative 34; Mismatches 94; Indels 3; Gaps 3;

Oy 1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
Db 1 MSKNDORTAKTEKTDWTLNHLFISSGLYKLNLSVAQITSLIAMIISTSLIAII 60
Oy 61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
Db 61 FIASANHKVLTTLTAIIODATSOIKNTPTTYLQDPOLGISFNSLSEITSOITTLASTTP 120
Oy 121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
Db 121 GVKSNLOPTVTKNTTQTQPSKPTTKORONKPNKPNNDHFVEVNFVPCISCSNMP 180
Oy 181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 239
181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 239
Db 181 TCMATCKRIPNKKPGKKTTPKPKPTTKKDKLPQTTPKREVTTPKEEPTINTK 239
Oy 240 KTNITITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 292
240 KTNITITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 292
Db 240 KTNITITLLTNNTGNPKLTSGMETPHSTSEGNLSPSOVSTTSEHPSPSPPTTR 292

RESULT 11
MG218
major surface glycoprotein G - human respiratory syncytial virus (strain 18537)
C:Species: human respiratory syncytial virus
C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999
C:Accession: B32703
R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.
Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987
A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and
A:Reference number: B32703; MUID:87289657; PMID:2441388
A:Accession: B32703
A:Molecule type: mRNA
A:Residues: 1-292 <JOH>
A:Cross-references: GB:M7213; NID:g333942; PIDN:AAA47412.1; PID:g333943
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:41-63/Domain: transmembrane #status predicted <TMN>
F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 49.3%; Score 761; DB 1; Length 292;

Best Local Similarity 54.9%; Pred. No. 1.7e-42;
Matches 158; Conservative 35; Mismatches 93; Indels 2; Gaps 2;

QY 1 MSNKKDQRTAKTEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMITSLITAI 60
Db 1 MSNKKDQRTAKTEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMITSLITAI 60

QY 61 FIASANKVLTLLAIODATSOIKNTPTLYLTODPOLGISFNSLSEITTSQTTILASTP 120
Db 61 FIASANKVLTLLAIODATSOIKNTPTLYLTODPOLGISFNSLSEITTSQTTILASTP 120

QY 121 GYKSNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNNDPHEVNFVPCSTCSNP 180
Db 121 NTKSETHHTTAQTGKRITTTSTQTNKPSSTKSRGNPKPEKPDYHFEVNFVPCSTCSNP 180

QY 181 TCMAICKRIPNKKPGKTTTKPTKKPTTKT-KKDLKPDQTKPKPEPTTKPPEPTINT 239
Db 181 LCKSICKTIPSNKKPKKPIKPTTKPTTKT-KKDLKPDQTKPKPEPTTKPPEPTINT 239

QY 240 KTNITTTLLTNTGNPKLTSQMETFHSSTSGNLSPSQVSTTSPPTRQ 287
Db 241 ERDTSISOSTVLDITTPKYTIQOOSLHSTPSSTPSSTQIPTASE-PS 287

RESULT 12

nuclcapsid protein (version 2) - human respiratory syncytial virus
C:Species: human respiratory syncytial virus
C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999
C:Accession: A04026
R:Elango, N.; Venkatesan, S.
Nucleic Acids Res. 11, 5941-5951, 1983
A:Title: Amino acid sequence of respiratory syncytial virus capsid protein.
A:Reference number: A04026; MUID:83299261; PMID:6310521
A:Accession: A04026
A:Molecule type: mRNA
A:Residues: 1-467 <ELAN>
A:Cross-references: GB:X00001; MID:961215; PIDN:CAA24906.1; PID:961216
C:Genetics:
A:Gene: N
C:Superfamily: respiratory syncytial virus nucleocapsid protein
C:Keywords: nucleocapsid

Query Match 27.5%; Score 425; DB 1; Length 467;
Best Local Similarity 90.2%; Pred. No. 1.7e-20;
Matches 83; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 207 TFKTKKDLKPDQTKPKPEPTTKPTTKNTTNTTNTGNPKLTSQMETFH 266
Db 376 TFKTKKDLKPDQTKPKPEPTTKPTTKNTTNTTNTGNPKLTSQMETFH 266

QY 267 STSSEGNLSPSOVSTTSSEHPSQSPSPPTTRQ 298
Db 436 STSSEGNLSPSOVSTTSSEHPSQSPSPPTTRQ 467

RESULT 13

glycoprotein G - bovine respiratory syncytial virus (isolate FS-1) (fragment)
C:Species: bovine respiratory syncytial virus
C:Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
C:Accession: P00768
R:Mailpedit, S.K.; Samal, S.K.
J. Gen. Virol. 74, 2001-2004, 1993
A:Title: Sequence variability of the glycoprotein gene of bovine respiratory syncytial
A:Reference number: J02284; MUID:93389461; PMID:8376974
A:Accession: P00768
A:Molecule type: mRNA
A:Residues: 1-250 <MAL>
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:1-31/Domain: intracellular #status predicted <INT>
F:32-59/Domain: transmembrane #status predicted <TMN>

F:60-250/Domain: extracellular #status predicted <EXT>
F:78,156,226,241,244/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 2; Length 250;
Best Local Similarity 30.3%; Pred. No. 1.3e-12;
Matches 91; Conservative 35; Mismatches 104; Indels 70; Gaps 9;

QY 11 KTEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMITSLITAIIFIASANKVT 70
Db 4 KTLKAMKASKYFTVGLSCLYKFNKSLVQTAISTLMITLTSVLTAIYISGNKAK 63

QY 71 LTTAIODATSOIKNTPTLYLTODPOLGISFNSLSEITTSQTTILASTP 130
Db 64 PTKSTIQOQTQQPNHNSPFTENY-----KTHNSIOSTTISQILNDITTR 111

QY 131 VKTKNTTTOTOPSK-----PTTKQ-----RONKPPNNDPHEVNFVPCSTCSN 178
Db 112 GTTGHSTDETOSEKRIKSOSTLPTTRKPPINPSSESNPDHNNNSOTLPPVPCSTCEG 171

QY 179 NPTCAICKRIPNKKPGKTTTKPTKKPTTKKDLKPDQTKPKPEPTTKPPEPTINT 238
Db 172 NLACLSIQVGPGR-----APSRAPTI-TLKTKTKPTTK-----KP----- 207

QY 239 TKNTTTLLTNTGNPKLTSQMETFHSSTSGNLSPSQVSTTSPPTRQ 298
Db 208 -----IKTTI-----HHRTSPEAKIQPN-NIAAPQGISSPENHTNQ 245

RESULT 14

MGNZBR
major surface glycoprotein G - bovine respiratory syncytial virus (strain 391-2)
N:Alternate names: attachment glycoprotein G
C:Species: bovine respiratory syncytial virus
C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: A36408
R:Jerich, R.A.; Anderson, K.; Wertz, G.W.
J. Virol. 64, 5559-5569, 1990
A:Title: Nucleotide sequence analysis and expression from recombinant vectors demonstr
y syncytial virus.
A:Reference number: A36408; MUID:91012801; PMID:2214024
A:Accession: A36408
A:Molecule type: mRNA
A:Residues: 1-257 <LEA>
A:Cross-references: GB:M58307; MID:9210830; PIDN:AAA42810.1; PID:9210831
C:Genetics:
A:Gene: G
C:Superfamily: respiratory syncytial virus major surface glycoprotein G
C:Keywords: glycoprotein; transmembrane protein
F:45-62/Domain: transmembrane #status predicted <TMN>
F:3,85,127,149,233,251/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.4%; Score 299; DB 1; Length 257;
Best Local Similarity 31.6%; Pred. No. 1.4e-12;
Matches 92; Conservative 33; Mismatches 120; Indels 46; Gaps 8;

QY 1 MSNKKDQRTAKTEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMITSLITAI 60
Db 1 MSNKKDQRTAKTEKTDWTLNHLFISSGLYKLNLSVAQITLSILAMITSLITAI 60

QY 61 FIASANKVLTLLAIODATSOIKNTPTLYLTODPOLGISFNSLSEITTSQTTILASTP 120
Db 61 FIASANKVLTLLAIODATSOIKNTPTLYLTODPOLGISFNSLSEITTSQTTILASTP 120

QY 121 GYKSNLOPTVTKNTTTOTOPSK---PTTKORONKPPNKNNDP-----HEVNF 168
Db 109 SOLINTDITGTYGISTNETONKRIKSGSTLPATKPPINPSGSIPEPNHODHNNFOL 168

QY 169 NFVPCSTCSNPTCAICKRIPNKKPGKTTTKPTKKPTTKKDLKPDQTKPKPEPTT 228
Db 169 PYPVPCSTCEGNLACLSL-----IETRAPSRAPTI-TLKTKPKPKPTTK----- 212

QY 229 KPPEPTINTKNTTNTTLLTNTTGNPKLTSQMETFHSSTSGNLSPSQV 279
Db 229 KPPEPTINTKNTTNTTLLTNTTGNPKLTSQMETFHSSTSGNLSPSQV 279

Db 213 KPTK--TTIHHRTSPETKLOPKNNTATP---QOGILSSTEHTNQSTQI 257

RESULT 15

glycoprotein G - bovine respiratory syncytial virus (isolate A51908)

C;Species: bovine respiratory syncytial virus

```
C;Date: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999
```

C;Accession: JQ2284

R; Mallipedi, S.K.; Samal, S.K.
T Gen Virol 74: 2001-2004

A:Title: Sequence variability of E

A:Reference number: JQ2284; MUID:93389461; PMID:8376974

A;Accession: JQ2284

A;Molecule type: mR

A;Residues: 1-263 <MA

A; Experimental source:

C; keyword: cytoskeleton
C; keyword: respiratory syncytial virus
C; keyword: transmembrane protein

C;keywords: glycoprotein; transmembrane protein
E:1-38/Domain: intracellular #status predicted <CYT>

F:	39-66/Domain:	transmembrane	#status predicted
E:	1-38/Domain:	intracellular	nbcdas predicted

```
F;67-263/Domain: extracellular #status predicted <EX
```

F;127,163,251/Binding site: carbohydrate (Asn) (coval

[illegible]

Query Match	19.18;	Score 295.5;	DB 2;	Length 263;
Post Local similarity	31.79;	Pred NO 3	3e-12;	.

Best Local
Matches

Category	Count	Percentage
Malicious	80	33%
Conservative	33	14%
Misadventures	120	51%
Others	27	11%

QY 1 M S K N K D Q R T A K T L E K T W D T L N H L L F I S S G L Y K L N L K S V A Q I T L S I L A M I I S T S L I T A I I 60

[illegible]

Db 1 MSNHTHPKFTLKRAWKASKYFIVGLSCLYKFNLSVLQTAATSLAMITLTSLVITAI 600

C:\PROGRAM FILES\COMMON-FILES\MODEL-CITCENET\CITEMCOMETTI ASCEND 170

61 FIASANHKVTLTIAIQDATSIKNTPTLYLQDPQLGISFSNLSELTISQIWIILASITP 120

61 YTSVGNAAKAPTSKPTTOOTOPONHTPI.PTEHNHKSHTSTOSTTISOPNI--DITS 118

DB 01 11SVGNHNNF13NF11QQ1QQFQNNH1FDDF1EHNHNS1HH1Q1Q11EQQ1NA 2110 11

QY 121 GVK-----SNLOPTTVKTKNT-TTQOTQPSKPTTKORONKRPKNKNDEHFEVENFVPCS 174

[illegible]

Db 119 GTTYGHPINRTQNRKIKSQSTPLATRKLPINPL---ESNPENHQDHNSQTLPVPCS 174

1 7 5 T O C W W M C U A T C W P I D N W W D C V W W M T W K W K D I W B O T T K B 3 2 3

Search completed: May 1, 2003, 13:47:16
Job time : 21.6792 secs

Job time : 21.6792 secs


```

QY 1 MSKNDQRTAKTLEKWTWDTLNHLFLISSGLYKLNKSVQITLSILAMIISTLITAI 60
Db 1 MSKNDQRTAKTLEKWTWDTLNHLFLISSGLYKLNKSVQITLSILAMIISTLITAI 60
QY 61 FIASANHKVTLTALIODATSOIKNTPTTYLTODPOLGISFSNLSSEITSOITTLASTP 120
Db 61 FIASANHKVTLTALIODATSOIKNTPTTYLTODPOLGISFSNLSSEITSOITTLASTP 120
QY 121 GVSANLOPTVTKNTTQTOPSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180
Db 121 GVSANLOPTVTKNTTQTOPSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180
QY 181 TCWAICKRIPNKKRPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 240
Db 181 TCWAICKRIPNKKRPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 240
QY 241 TMTITLLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTSEHSPSSPPTTRQ 298
Db 241 TMTITLLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTSEHSPSSPPTTRQ 298

RESULT 2
VGLG_HRSVA STANDARD: PRT: 298 AA.
ID VGLG_HRSVA STANDARD: PRT: 298 AA.
AC P03423;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain A2).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11259;
RN [1]
RP SEQUENCE FROM N.A.
RA Wertz G.W., Collins P.L., Huang Y., Gruber C., Levine S., Ball L.A.;
RT "Nucleotide sequence of the G protein gene of human respiratory
RT syncytial virus reveals an unusual type of viral membrane protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4075-4079(1985).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=86067198; PubMed=4069997;
RA Satake M., Colligan J.E., Elango N., Norrby E., Venkatesan S.;
RT "Respiratory syncytial virus envelope glycoprotein (G) has a novel
RT structure".
RL Nucleic Acids Res. 13:7795-7812(1985).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=95265253; PubMed=7747420;
RA Connors M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;
RT "A cold-passaged, attenuated strain of human respiratory syncytial
RT virus contains mutations in the F and L genes.";
RL Virology 208:478-484(1995).
RN [4]
RP SEQUENCE FROM N.A.
RA MEDLINE=97187925; PubMed=9035372;
RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,
RA Murphy B.R.;
RT "Acquisition of the ts phenotype by a chemically mutagenized cold-
RT passaged human respiratory syncytial virus vaccine candidate results
RT from the acquisition of a single mutation in the polymerase (L)
RT gene.";
RL Virus Genes 13:269-273(1996).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.

```

```

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M11486; AAB59857.1; -
DR EMBL: X03149; CA26928.1; -
DR EMBL: U05062; AAB86663.1; -
DR EMBL: U05063; AAB86675.1; -
DR EMBL: U06364; AAC55969.1; -
DR PIR: A04039; MGZ.
DR InterPro: IPR000925; Glycoprotein_G.
DR Pfam: PF00802; Glycoprotein_G.1.
KM Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66
FT DOMAIN 67 298
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 298 AA; 32586 MW; 993C3D2DD68BC634 CRC64;

Query Match 91.8%; Score 1418; DB 1; Length 298;
Best Local Similarity 93.3%; Pred. No. 3e-86;
Matches 278; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 MSKNDQRTAKTLEKWTWDTLNHLFLISSGLYKLNKSVQITLSILAMIISTLITAI 60
Db 1 MSKNDQRTAKTLEKWTWDTLNHLFLISSGLYKLNKSVQITLSILAMIISTLITAI 60
QY 61 FIASANHKVTLTALIODATSOIKNTPTTYLTODPOLGISFSNLSSEITSOITTLASTP 120
Db 61 FIASANHKVTLTALIODATSOIKNTPTTYLTODPOLGISFSNLSSEITSOITTLASTP 120
QY 121 GVSANLOPTVTKNTTQTOPSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180
Db 121 GVSANLOPTVTKNTTQTOPSKPTKORONKPPNKNPNNDHFVFNVPSCISNMP 180
QY 181 TCWAICKRIPNKKRPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 240
Db 181 TCWAICKRIPNKKRPGKKTTPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 240
QY 241 TMTITLLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTSEHSPSSPPTTRQ 298
Db 241 TMTITLLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTSEHSPSSPPTTRQ 298

RESULT 3
VGLG_HRSV3 STANDARD: PRT: 297 AA.
ID VGLG_HRSV3 STANDARD: PRT: 297 AA.
AC P27022;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb1734).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11253;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND

```



```

CC      HEMAGGLUTININATING ACTIVITIES.
CC      -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC      CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC      DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR      PIR: J01205; J01205.
DR      InterPro: IPR000925; Glycoprot_G.
DR      Pfam: PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.
FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 297
FT      CARBOHYD 135 135
FT      CARBOHYD 237 237
FT      CARBOHYD 251 251
SQ      SEQUENCE 297 AA; 32525 MW; 48448F9E091E1802 CRC64;

Query Match      87.0%; Score 1343; DB 1; Length 297;
Best Local Similarity 88.6%; Pred. No. 2,4e-81;
Matches 263; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY      1 MSKNDQRTAKTLEKTPDNLNHLFISSGLYKLNKSAVOITSLIAMIISTLIITAI 60
DB      1 MSKNDQRTAKTLEKTPDNLNHLFISSGLYKLNKSAVOITSLIAMIISTLIITAI 60
QY      61 FIASANKHVLTITAIIDATSIQIKNTTPYLTQDPOLGISFSNLSEITSQTTILASTTP 120
DB      61 FIASANKHVLTITAIIDATSIQIKNTTPYLTQDPOLGISFSNLSEITSQTTILASTTP 120
QY      121 GYKSNLOPTTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
DB      121 GYKSNLOPTTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
QY      121 GYKSNLOPTTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
DB      121 GYKSNLOPTTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
QY      181 TCMAICKRIPNKKPKGKTTTKPKPKPTTKKDLKPOTTKPKPEVPTTKPTERTINTTK 240
DB      181 TCMAICKRIPNKKPKGKTTTKPKPKPTTKKDLKPOTTKPKPEVPTTKPTERTINTTK 240
QY      181 TCMAICKRIPNKKPKGKTTTKPKPKPTTKKDLKPOTTKPKPEVPTTKPTERTINTTK 240
DB      181 TCMAICKRIPNKKPKGKTTTKPKPKPTTKKDLKPOTTKPKPEVPTTKPTERTINTTK 240
QY      241 TNITTTLLTNNTGNPKLTISOMETHSTSEGNLSPSOVSTTSEHPSPSSPNTT 297
DB      241 TNITTTLLTNNTGNPKLTISOMETHSTSEGNLSPSOVSTTSEHPSPSSPNTT 297

RESULT 4
VGLG_HRSV6      STANDARD:      PRT:      297 AA.
ID      VGLG_HRSV6
AC      P27025;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      01-AUG-1992 (Rel. 23, Last annotation update)
DE      Major surface glycoprotein G (Attachment glycoprotein G).
GN      G.
OS      Human respiratory syncytial virus (strain rsb6256).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_TaxID=11256;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91374005; PubMed=1895054;
RA      Cane P.A., Matthews D.A., Pringle C.R.;
RT      "Identification of variable domains of the attachment (G) protein of
RT      subgroup A respiratory syncytial viruses.";
RL      J. Gen. Virol. 72:2091-2096(1991).
CC      -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC      RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC      HEMAGGLUTININATING ACTIVITIES.
CC      -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC      CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC      DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR      PIR: J01208; J01208.
DR      InterPro: IPR000925; Glycoprot_G.
DR      Pfam: PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.

```

```

FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 297
FT      CARBOHYD 103 103
FT      CARBOHYD 135 135
FT      CARBOHYD 237 237
FT      CARBOHYD 251 251
SQ      SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

Query Match      87.0%; Score 1343; DB 1; Length 297;
Best Local Similarity 88.3%; Pred. No. 2,4e-81;
Matches 262; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

QY      1 MSKNDQRTAKTLEKTPDNLNHLFISSGLYKLNKSAVOITSLIAMIISTLIITAI 60
DB      1 MSKNDQRTAKTLEKTPDNLNHLFISSGLYKLNKSAVOITSLIAMIISTLIITAI 60
QY      61 FIASANKHVLTITAIIDATSIQIKNTTPYLTQDPOLGISFSNLSEITSQTTILASTTP 120
DB      61 FIASANKHVLTITAIIDATSIQIKNTTPYLTQDPOLGISFSNLSEITSQTTILASTTP 120
QY      121 GYKSNLOPTTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
DB      121 GYKSNLOPTTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
QY      121 SAESTPOSTVTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
DB      121 SAESTPOSTVTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVNFVPCISGNNP 180
QY      181 TCMAICKRIPNKKPKGKTTTKPKPKPTTKKDLKPOTTKPKPEVPTTKPTERTINTTK 240
DB      181 TCMAICKRIPNKKPKGKTTTKPKPKPTTKKDLKPOTTKPKPEVPTTKPTERTINTTK 240
QY      241 TNITTTLLTNNTGNPKLTISOMETHSTSEGNLSPSOVSTTSEHPSPSSPNTT 296
DB      241 TNITTTLLTNNTGNPKLTISOMETHSTSEGNLSPSOVSTTSEHPSPSSPNTT 296

RESULT 5
VGLG_HRSV7      STANDARD:      PRT:      297 AA.
ID      VGLG_HRSV7
AC      P27026;
DT      01-AUG-1992 (Rel. 23, Created)
DT      01-AUG-1992 (Rel. 23, Last sequence update)
DT      01-AUG-1992 (Rel. 23, Last annotation update)
DE      Major surface glycoprotein G (Attachment glycoprotein G).
GN      G.
OS      Human respiratory syncytial virus (strain rsb6614).
OC      Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC      Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX      NCBI_TaxID=11257;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91374005; PubMed=1895054;
RA      Cane P.A., Matthews D.A., Pringle C.R.;
RT      "Identification of variable domains of the attachment (G) protein of
RT      subgroup A respiratory syncytial viruses.";
RL      J. Gen. Virol. 72:2091-2096(1991).
CC      -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC      RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC      HEMAGGLUTININATING ACTIVITIES.
CC      -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC      CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC      -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC      DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR      PIR: J01209; J01209.
DR      InterPro: IPR000925; Glycoprot_G.
DR      Pfam: PF00802; Glycoprotein_G; 1.
KW      Transmembrane; Glycoprotein.
FT      DOMAIN 1 37
FT      TRANSMEM 38 66
FT      DOMAIN 67 297
FT      CARBOHYD 103 103
FT      CARBOHYD 135 135
FT      CARBOHYD 237 237
FT      CARBOHYD 251 251
SQ      SEQUENCE 297 AA; 32708 MW; 6781756C38B64A80 CRC64;

```

FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA: 32670 MW: 588384028E437ACD CRC64;
Query Match 86.0%; Score 1328; DB 1; Length 297;
Best Local Similarity 87.8%; Pred. No. 2.3e-80;
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60
DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSIAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTFTAIIDQATSQIKNTPTVLTQDPOLGIFSNLSSETSTTTTILASTP 120
DB 61 FIASANHKVLTFTAIIDQATSQIKNTPTVLTQDPOLGIFSNLSSETSTPTTPALTTP 120
QY 121 GKSNLOPTVTKNTTGTOTOPSKPTTKORONKPPNNDPHEVFNFPVPCISCSNNP 180
DB 121 SAKSTPSTVTKNTTGTOTOPSKPTTKORONKPPNNDPHEVFNFPVPCISCSNNP 180
QY 181 TCWAICRIPNKKPKKTK 240
DB 181 TCWAICRIPNKKPKKTK 240
QY 241 TMTITLLTNTGNTGNPKLTSMETPHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296
DB 241 TMTITLLTNTGNTGNPKLTSMETPHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296

RESULT 6
VGLG_HRSV4 STANDARD; PRT; 297 AA.
AC P27023;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN Human respiratory syncytial virus (strain rsb5857).
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11254;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR: J01206; J01206.
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 297 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 297 AA: 32772 MW: 104886CA475936BE CRC64;

Query Match 85.8%; Score 1324; DB 1; Length 297;
Best Local Similarity 87.8%; Pred. No. 4.2e-80;
Matches 260; Conservative 7; Mismatches 29; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60
DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSIAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTFTAIIDQATSQIKNTPTVLTQDPOLGIFSNLSSETSTTTTILASTP 120
DB 61 FIASANHKVLTFTAIIDQATSQIKNTPTVLTQDPOLGIFSNLSSETSTPTTPALTTP 120
QY 121 GKSNLOPTVTKNTTGTOTOPSKPTTKORONKPPNNDPHEVFNFPVPCISCSNNP 180
DB 121 SAKSTPSTVTKNTTGTOTOPSKPTTKORONKPPNNDPHEVFNFPVPCISCSNNP 180
QY 181 TCWAICRIPNKKPKKTK 240
DB 181 TCWAICRIPNKKPKKTK 240
QY 241 TMTITLLTNTGNTGNPKLTSMETPHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296
DB 241 TMTITLLTNTGNTGNPKLTSMETPHSSSEGNLSPSOVSTTSEHPSQSPSPNTT 296

RESULT 7
VGLG_HRSV5 STANDARD; PRT; 298 AA.
AC P27024;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN Human respiratory syncytial virus (strain rsb190).
OS Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11255;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR: J01207; J01207.
DR InterPro: IPR000925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 67 298 POTENTIAL.
FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 298 AA: 32769 MW: 4D74E854D34D7B85 CRC64;

Query Match 85.0%; Score 1313; DB 1; Length 298;
Best Local Similarity 86.9%; Pred. No. 2.2e-79;
Matches 259; Conservative 11; Mismatches 28; Indels 0; Gaps 0;

QY 1 MSKNKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSVAQITLSILAMISTSLITAI 60
DB 1 MSKTKDQRTAKTLEKTDWTLNHLFISSGLYKLNKLSIAQITLSILAMISTSLITAI 60
QY 61 FIASANHKVLTFTAIIDQATSQIKNTPTVLTQDPOLGIFSNLSSETSTTTTILASTP 120
DB 61 FIASANHKVLTFTAIIDQATSQIKNTPTVLTQDPOLGIFSNLSSETSTPTTPALTTP 120

[illegible]

```

Db 1 MSKHKNQGFATLEKTMWDLNLHLIVISSCLYRLNKLSTIAQLATLVLMISTSLIAAI 60
QY 61 FIASANKVLTTLTALIODATSOIKNTPTPTTYLTQDPOLGISNSLHITSQTTLIASTTP 120
Db 61 FIASANKVLTTLTALIODATSOIKNTPTPTTYLTQDPOLGISNSLHITSQTTLIASTTP 120
QY 121 GVKSLAQFTTKTKNTNTTQTOPSPKPTKQKONKPPNKPNNDPFEVEFNVCISCSNP 180
Db 121 NTKSSTHTHTAQVGRITTSQTQNKPSKSKSNKPKPKRDYHFEVFNFPSCICGNG 180
QY 181 TCMAICKRIPNKKPGKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 239
Db 181 LCKSLCKITPSNKKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPKPK 240
QY 240 KTNITTTLLTNNTGNNPKLTSOMEFHSTSSSGNLSPOVSTTSEHPS 287
Db 241 ERDTISQSTVLDITTPKTYITQOQSLHSTSTSENPSTSTPTASE-PS 287

RESULT 11
VGLG_BRSV1 STANDARD; PRT; 263 AA.
ID VGLG_BRSV1
AC 010683;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain 127) (BRV).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=82818;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97286324; PubMed=9143302;
RA Furze J., Roberts S., Wertz G., Taylor G.;
RT "Antigenically distinct G glycoproteins of BRV strains share a high
RL degree of genetic homogeneity.";
RL Virology 231:48-58(1997).
CC -I- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -I- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -I- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
CC HRS VIRUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: Y08716; CAA69966.1; -.
DR HSPF: P2261; IBRV.
DR InterPro: IPR00925; Glycoprot_G.
DR Pfam: PF00802; Glycoprotein_G. 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66 POTENTIAL.
FT DOMAIN 67 263 EXTRACELLULAR (POTENTIAL).
FT DISULFID 173 186 BY SIMILARITY.
FT DISULFID 176 182 BY SIMILARITY.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 263 AA; 28948 MW; D592D7889899535 CRC64;

```

```

QY 1 MSKNKQDRTAKTLEKTDLTUNHLFISSGLYKLNKSVAOITSLIAMIISLITAI 60
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MSNHTHHPRFKTLKRAKMAKSKFYIVGLSCLYKFNLSVOTALTTLAMITLTSVITAI 60
QY 61 FIASANKVTLTITAIIDATSOIKNTTPYLYLODPOLGISFSNLSSETTSQTTILASTP 120
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 YISVGNAKAKPTSKPTTQOTLOLQNHPTPLLEHNY-----KSTHTSISQSTTL 108
QY 121 GYKSNLQPTTYKTKNTTTTQTOPSK---PTTKRONKPPKPP-----NNDHF--FEVF 168
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 109 SOPNIDTTSSTTYGHPPTNRTONRKIKSQSTPLATRKPPINPLGSPNPNHODHNSQTL 168
QY 169 NFVPCSSICNNPTCWAIC---KRIPN-----KKPKKTTTKPTKPPFKTKDKDL 216
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 169 PHVPCSTCEGNPCSPLOQIELEKAPSSAPITLLKAKPKTKTKPTTIIHRTISPEAK 228
QY 217 POTTREKVEPTTKPTPEPTINTTKNTITTLTNNT 253
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 229 LOTKKIMATPOOGLILSP---EHQTNOSTTOISQHTS 262

RESULT 12
VGLG_BRSV4 STANDARD: PRT; 263 AA.
AC 010684;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain 4642) (BRS).
OC Paramyxoviridae; ssRNA negative-strand viruses; Mononegavirales;
OX NCBI_TaxID=82820;
RN [1]
RP SEQUENCE FROM N. A.
RA MEDLINE-97286324; PubMed-9143302;
RA Furze J., Roberts S., Wertz G., Taylor G.;
RT "Antigenically distinct G glycoproteins of BRSV strains share a high
  degree of genetic homogeneity.";
RL Virology 231:48-58(1997)
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
  RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
  HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
  CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
  HRS VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y08718; CAA69968.1; -
DR HSSP; P22261; IRRV.
DR InterPro; IPR000925; Glycoprotei_G.
DR Pfam; PF00802; Glycoprotein_G.1.
KM Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 38 66 POTENTIAL.
FT DOMAIN 67 263 EXTRACELLULAR (POTENTIAL).
FT DISULFID 173 186 BY SIMILARITY.
FT DISULFID 176 182 BY SIMILARITY.
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 163 163 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 263 AA; 28976 MW; D592D79EF8EE9535 CRC64;

```

```

Query Match 20.0%; Score 308.5; DB 1; Length 263;
Best Local Similarity 32.5%; Pred. No. 9.0e-14;
Matches 90; Conservative 32; Mismatches 116; Indels 39; Gaps 7;

QY 1 MSKNKQDRTAKTLEKTDLTUNHLFISSGLYKLNKSVAOITSLIAMIISLITAI 60
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MSNHTHHPRFKTLKRAKMAKSKFYIVGLSCLYKFNLSVOTALTTLAMITLTSVITAI 60
QY 61 FIASANKVTLTITAIIDATSOIKNTTPYLYLODPOLGISFSNLSSETTSQTTILASTP 120
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 YISVGNAKAKPTSKPTTQOTLOLQNHPTPLLEHNY-----KSTHTSISQSTTL 108
QY 121 GYKSNLQPTTYKTKNTTTTQTOPSK---PTTKRONKPPKPP-----NNDHF--FEVF 168
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 109 SOPNIDTTSSTTYGHPPTNRTONRKIKSQSTPLATRKPPINPLGSPNPNHODHNSQTL 168
QY 169 NFVPCSSICNNPTCWAIC---KRIPN-----KKPKKTTTKPTKPPFKTKDKDL 216
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 169 PHVPCSTCEGNPCSPLOQIELEKAPSSAPITLLKAKPKTKTKPTTIIHRTISPEAK 228
QY 217 POTTREKVEPTTKPTPEPTINTTKNTITTLTNNT 253
  |||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 229 LOTKKIMATPOOGLILSP---EHQTNOSTTOISQHTS 262

RESULT 13
VGLG_BRSV4 STANDARD: PRT; 257 AA.
AC P22261;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Bovine respiratory syncytial virus (strain Copenhagen) (BRS).
OC Paramyxoviridae; ssRNA negative-strand viruses; Mononegavirales;
OX NCBI_TaxID=11248;
RN [1]
RP SEQUENCE FROM N. A.
RA MEDLINE-91012801; PubMed-2214024;
RA Lerch R.A., Anderson K., Wertz G.W.;
RT "Nucleotide sequence analysis and expression from recombinant vectors
  demonstrate that the attachment protein G of bovine respiratory
  syncytial virus is distinct from that of human respiratory syncytial
  virus.";
RL J. Virol. 64:5559-5569(1990).
RN [2]
RP STRUCTURE BY NMR OF 158-189.
RC STRAIN-391-2;
RX MEDLINE-97098087; PubMed-8942628;
RA Dorelmeijers J.F., Langendijk J.P.M., Haard K., Boelens R.,
  Rullmann J.A., Schaeper W.M., van Oirschot J.T., Kaptein R.;
RT "Solution structure of the immunodominant region of protein G of
  bovine respiratory syncytial virus.";
RL Biochemistry 35:14684-14688(1996).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
  RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
  HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
  CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- SIMILARITY: TO THE G PROTEINS OF EITHER THE SUBGROUP A OR B
  HRS VIRUS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M58307; AAA42810.1; -

```


GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:27:43 ; Search time 35.9849 Seconds
(without alignments)
1706.328 Million cell updates/sec

Title: US-09-462-816-2

Perfect score: 1544

Sequence: 1 MSKNKDORTAFLEKTDLT.....VSTTSEHPSPSPPTNTTQ 298

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL.21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_prodent:*
12: sp_virus:*
13: sp Vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1537	99.5	298	12	001929 human respi
2	1533	99.3	298	12	082068 human respi
3	1393	90.2	298	12	009634 human respi
4	1389	90.0	298	12	009719 respiratory
5	1365	88.4	297	12	09YVB5 human respi
6	1364	88.3	297	12	082066 human respi
7	1356	87.8	298	12	09YVB4 human respi
8	1354	87.7	297	12	082071 human respi
9	1345	87.1	297	12	082067 human respi
10	1344	87.0	297	12	082058 human respi
11	1339	86.7	297	12	091947 human respi
12	1336	86.5	293	12	09YVB3 human respi
13	1334	86.4	295	12	086359 respiratory
14	1334	86.4	295	12	082065 respiratory
15	1334	86.4	298	12	082065 human respi
16	1331	86.2	297	12	082074 human respi

17	1331	86.2	297	12	082064 human respi
18	1328	86.0	297	12	082063 human respi
19	1326	85.9	292	12	09YVB2 human respi
20	1326	85.9	299	12	082077 human respi
21	1325	85.8	298	12	082060 human respi
22	1324	85.8	297	12	091946 human respi
23	1320	85.5	297	12	082070 human respi
24	1319	85.4	293	12	09YVB1 human respi
25	1319	85.4	295	12	086361 respiratory
26	1319	85.4	297	12	082079 human respi
27	1319	85.4	297	12	082057 human respi
28	1318	85.4	297	12	09YVC8 human respi
29	1317	85.3	298	12	082075 human respi
30	1317	85.3	298	12	082062 human respi
31	1317	85.3	298	12	082056 human respi
32	1316	85.2	297	12	082072 human respi
33	1315	85.2	295	12	086357 respiratory
34	1314	85.1	297	12	082078 human respi
35	1314	85.1	297	12	091944 human respi
36	1314	85.1	298	12	082073 human respi
37	1314	85.1	298	12	082076 human respi
38	1312	85.0	298	12	011978 human respi
39	1311	84.9	292	12	09YVB0 human respi
40	1311	84.9	298	12	091945 human respi
41	1308	84.7	296	12	086358 respiratory
42	1306	84.6	298	12	082069 human respi
43	1305	84.5	298	12	082061 human respi
44	1298	84.1	295	12	086356 respiratory
45	1282	83.0	279	12	090678 human respi

ALIGNMENTS

RESULT 1
ID 001929 PRELIMINARY; PRT; 298 AA.
AC 001929;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup A / strain Long).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11260;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-9106351; PubMed-2249671;
RA Garcia-Barreno B., Portela A., Delgado T., Lopez J.A., Melero J.A.;
RT "Frame shift mutations as a novel mechanism for the generation of
RT neutralization resistant mutants of human respiratory syncytial
RT virus.";
RL EMBL J. 9:4181-4187(1990).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: THIS PROTEIN MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE
CC CHAINS DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC EMBL: X17085; CAA34937.1; -;
DR InterPro: IPR000925; Glycoprot G.
DR InterPro: IPR003880; Prntine.attach.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOAMPHETHEINE; UNKNOWN_1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 INTRACELLULAR (BY SIMILARITY).
FT TRANSMEM 38 66
FT DOMAIN 67 298 EXTRACELLULAR (BY SIMILARITY).
FT DOMAIN 205 211 ESSENTIAL FOR EPITOPE 636 INTEGRITY.
FT SITE 64 65 CLEAVAGE (BY PROTEASE) (POTENTIAL).

FT SITE 74 75 CLEAVAGE (BY PROTEASE) (POTENTIAL)
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 103 103 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL)
 FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL)
 SQ SEQUENCE 298 AA; 32771 MW; F8CD4213D97C2952 CRC64;

Query Match 99.5%; Score 1537; DB 12; Length 298;
 Best Local Similarity 99.3%; Pred. No. 4, 5e-101;
 Matches 296; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKLEKTWDTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
 DB 1 MSKNKQRTAKLEKTWDTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
 QY 61 FIASANHKVLTALIODATSOIKNTPTVLTODPOLGISFSNLSSETTTLASTP 120
 DB 61 FIASANHKVLTALIODATSOIKNTPTVLTODPOLGISFSNLSSETTTLASTP 120
 QY 121 GVSXNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNPNNDHFVFNVPSCISNNP 180
 DB 121 GVSXNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNPNNDHFVFNVPSCISNNP 180
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPTFKTKDLKQDTTKPREVPTKPEEPTINTTK 240
 DB 181 TCMAICKRIIPNKKPGKKTTPKPKPTFKTKDLKQDTTKPREVPTKPEEPTINTTK 240
 QY 241 TTTTLLTNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298
 DB 241 TTTTLLTNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298

RESULT 2

Q82068 PRELIMINARY; PRT; 298 AA.
 AC Q82068;
 DT 01-NOV-1996 (Tremblrel. 01, Created)
 DT 01-NOV-1996 (Tremblrel. 01, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE Glycoprotein (Fragment).
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBL_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SUBGROUP A;
 RX MEDLINE=9435057; PubMed=8057427;
 RA Garcia O., Martin M., Dopazo J., Arbliza J., Fabrasalle S., Russi J.,
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
 RA Melero J.A.;
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
 A): co-circulating lineages and correlation of genetic and antigenic
 changes in the G glycoprotein."
 RL J. VIROL. 68:3448-3459(1994).
 DR EMBL: Z33429; CAAB3872.1; -
 DR InterPro: IPR000925; Glycoprot.G.
 DR InterPro: IPR003880; Pantine.attach.
 DR Pfam: PF00802; Glycoprotein.G; 1
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 298 AA; 32753 MW; 7D080307897A772B CRC64;

Query Match 99.3%; Score 1533; DB 12; Length 298;
 Best Local Similarity 99.0%; Pred. No. 8, 6e-101;
 Matches 293; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSKNKQRTAKLEKTWDTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
 DB 1 MSKNKQRTAKLEKTWDTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60

QY 61 FIASANHKVLTALIODATSOIKNTPTVLTODPOLGISFSNLSSETTTLASTP 120
 DB 61 FIASANHKVLTALIODATSOIKNTPTVLTODPOLGISFSNLSSETTTLASTP 120
 QY 121 GVSXNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNPNNDHFVFNVPSCISNNP 180
 DB 121 GVSXNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNPNNDHFVFNVPSCISNNP 180
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPTFKTKDLKQDTTKPREVPTKPEEPTINTTK 240
 DB 181 TCMAICKRIIPNKKPGKKTTPKPKPTFKTKDLKQDTTKPREVPTKPEEPTINTTK 240
 QY 241 TTTTLLTNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298
 DB 241 TTTTLLTNNTGNPKLTQSOMETPHSTSEGNLSPSQVSTSEHSPSSPPTTRQ 298

RESULT 3

O09634 PRELIMINARY; PRT; 298 AA.
 ID O09634;
 AC O09634;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, last annotation update)
 DE Attachment protein (G).
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBL_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S2;
 RX MEDLINE=97185152; PubMed=9032893;
 RA Tolley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,
 RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
 RA Pringle C.R.;
 RT "Identification of mutations contributing to the reduced virulence of
 a modified strain of respiratory syncytial virus."
 RL Vaccine 14:1637-1646(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S2;
 RA Easton A.J.;
 RX Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U39662; AAC57026.1; -
 DR InterPro: IPR000925; Glycoprot.G.
 DR InterPro: IPR003880; Pantine.attach.
 DR Pfam: PF00802; Glycoprotein.G; 1
 DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D4997F31 CRC64;

Query Match 90.2%; Score 1393; DB 12; Length 298;
 Best Local Similarity 91.3%; Pred. No. 6, 5e-91;
 Matches 272; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 1 MSKNKQRTAKLEKTWDTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
 DB 1 MSKNKQRTAKLEKTWDTLNHLFISSGLYKLNKSAQITLSILAMISTSLITAI 60
 QY 61 FIASANHKVLTALIODATSOIKNTPTVLTODPOLGISFSNLSSETTTLASTP 120
 DB 61 FIASANHKVLTALIODATSOIKNTPTVLTODPOLGISFSNLSSETTTLASTP 120
 QY 121 GVSXNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNPNNDHFVFNVPSCISNNP 180
 DB 121 GVSXNLOPTVTKNTTTOTOPSKPTTKORONKPPNKNPNNDHFVFNVPSCISNNP 180
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPTFKTKDLKQDTTKPREVPTKPEEPTINTTK 240
 DB 181 TCMAICKRIIPNKKPGKKTTPKPKPTFKTKDLKQDTTKPREVPTKPEEPTINTTK 240

QY 241 TNITTTLLTNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSHHPQSSPPTTRQ 298
 Db 241 TNITTTLLTNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSHHPQSSPPTTRQ 298

RESULT 4

009719 PRELIMINARY; PRT; 298 AA.
 ID 009719;
 AC 009719;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Attachment glycoprotein (g).
 GN G.
 OS Respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxId=12814;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S2 TSIC;
 RX MEDLINE=97185152; PubMed=9032893;
 RA Tooley K.P., Marriott A.C., Simpson A., Plows D.J., Matthews D.A.,
 RA Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
 RA Pringle C.R.;
 RT "Identification of mutations contributing to the reduced virulence of
 RT a modified strain of respiratory syncytial virus.";
 RL Vaccine 14:1637-1646(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S2 TSIC;
 RL Easton A.J.;
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U39661; AAC57036.1; -;
 DR InterPro: IPR000925; Glycoprot_G.
 DR InterPro: IPR003880; Ppantne_attach.
 DR Pfam: PF00802; Glycoprotein_G; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 SQ SEQUENCE 298 AA; 32779 MW; 67FAA043682FAA50 CRC64;

Query Match 90.0%; Score 1389; DB 12; Length 298;
 Best Local Similarity 90.9%; Pred. No. 1.2e-90;
 Matches 271; Conservative 7; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKKNQRTAKTEKTEWDTLNHLFISSGLYKLNKSAQTSLIAMIISTLIITAI 60
 Db 1 MSKKNQRTAKTEKTEWDTLNHLFISSGLYKLNKSAQTSLIAMIISTLIITAI 60

QY 61 FIASANHKVLTITAIIDATISQIKNTPTPYLTQDPOLGISFNSLEITSOTTTLASTTP 120
 Db 61 FIASANHKVLTITAIIDATISQIKNTPTPYLTQDPOLGISFNSLEITSOTTTLASTTP 120

QY 121 GYKSNLOPTVTKKNTTTTQTOPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180
 Db 121 GYKSNLOPTVTKKNTTTTQTOPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180

QY 121 SVKSTLSTVTKTKNTTTTKIQPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180
 Db 121 SVKSTLSTVTKTKNTTTTKIQPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180

QY 181 TCMAICKRIPKKPGKRTTKPKPKPTFKTKDKDKPQTKPKKEVPTTKPEEPTINTTK 240
 Db 181 TCMAICKRIPKKPGKRTTKPKPKPTFKTKDKDKPQTKPKKEVPTTKPEEPTINTTK 240

QY 241 TNITTTLLTNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSHHPQSSPPTTRQ 298
 Db 241 TNITTTLLTNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSHHPQSSPPTTRQ 298

RESULT 5
 09YVB5 PRELIMINARY; PRT; 297 AA.
 ID 09YVB5;
 AC 09YVB5;
 DT 01-MAY-1999 (TReMBLrel. 10, Created)
 DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Attachment glycoprotein G.

OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxId=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WV2780;
 RX MEDLINE=99022964; PubMed=9806017;
 RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
 RT "Antigenic and genetic diversity among the attachment proteins of
 RT group A respiratory syncytial viruses that have caused repeat
 RT infections in children.";
 RL J. Infect. Dis. 178:925-932(1998).
 DR EMBL: AF065405; AAD02941.1; -;
 DR InterPro: IPR000925; Glycoprot_G.
 DR InterPro: IPR003880; Ppantne_attach.
 DR InterPro: IPR002965; P-rich_extensn.
 DR Pfam: PF00802; Glycoprotein_G; 1.
 DR PRINTS: PR01217; PRICEXTENS.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 SQ SEQUENCE 297 AA; 32572 MW; 5C10E719A05EF5C1 CRC64;

Query Match 88.4%; Score 1365; DB 12; Length 297;
 Best Local Similarity 90.2%; Pred. No. 6.1e-89;
 Matches 268; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSKKNQRTAKTEKTEWDTLNHLFISSGLYKLNKSAQTSLIAMIISTLIITAI 60
 Db 1 MSKKNQRTAKTEKTEWDTLNHLFISSGLYKLNKSAQTSLIAMIISTLIITAI 60

QY 61 FIASANHKVLTITAIIDATISQIKNTPTPYLTQDPOLGISFNSLEITSOTTTLASTTP 120
 Db 61 FIASANHKVLTITAIIDATISQIKNTPTPYLTQDPOLGISFNSLEITSOTTTLASTTP 120

QY 121 GYKSNLOPTVTKKNTTTTQTOPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180
 Db 121 GYKSNLOPTVTKKNTTTTQTOPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180

QY 121 SVKSTLSTVTKTKNTTTTKIQPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180
 Db 121 SVKSTLSTVTKTKNTTTTKIQPSKPTTKORONKPPKPNNDFFHFVFNPCGICSNP 180

QY 181 TCMAICKRIPKKPGKRTTKPKPKPTFKTKDKDKPQTKPKKEVPTTKPEEPTINTTK 240
 Db 181 TCMAICKRIPKKPGKRTTKPKPKPTFKTKDKDKPQTKPKKEVPTTKPEEPTINTTK 240

QY 241 TNITTTLLTNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSHHPQSSPPTTR 297
 Db 241 TNITTTLLTNNTGNPKLTSOMETFHSTSEGNLSPSQVSTTSHHPQSSPPTTR 297

RESULT 6
 082066 PRELIMINARY; PRT; 297 AA.
 ID 082066;
 AC 082066;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Glycoprotein.
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxId=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SUBGROUP A;
 RX MEDLINE=94335057; PubMed=8057427;
 RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabriasle S., Russi J.,
 RA Hortal M., Perez-Brena F., Martinez I., Garcia-Barreno B.,
 RA Melero J.A.;
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
 RT A): cocirculating lineages and correlation of genetic and antigenic
 RT changes in the G glycoprotein.";
 RL J. Virol. 68:5448-5459(1994).
 DR EMBL: Z33427; CAA83870.1; -;
 DR InterPro: IPR000925; Glycoprot_G.

DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00802; Glycoprotein_G; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 SO SEQUENCE 297 AA; 32589 MW; F210F6C68193F5DD CRC64;

Query Match 88.3%; Score 1364; DB 12; Length 297;
 Best Local Similarity 89.9%; Pred. No. 7.2e-89;
 Matches 267; Conservative 10; Mismatches 20; Indels 0; Gaps 0;

QY 1 MSKNKQDRTAKTLEKMTDNLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60
 Db 1 MSKNKQDRTAKTLEKMTDNLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60
 QY 61 FIASANHKVLTLLTAIIODATSOIKNTTPYVLTQDPOLGISFNSLSEITTSQTTIIASTTP 120
 Db 61 FIASANHKVLTLLTAIIODATNOIKNTTPYVLTQNPOLGISFNSPSEITSLITIIIDSTTP 120
 QY 121 GYKSNLOPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEVEFNVPSCISGNNP 180
 Db 121 GYKSTLOSTTVGKTNTTTQAOQPNKPTTKORONKPPNKNDFEVEFNVPSCISGNNP 180
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPTTKKDKLPQTTKPEVEPTTKPTEPTTINTTK 240
 Db 181 TCMAICKRIIPNKKPGKKTTPKPKPTTKKDKLPQTTKPEVEPTTKPTEPTTINTTK 240
 QY 241 TINTTLLTNNTGNPKLTSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTTR 297
 Db 241 TINTTLLTNTNRNPELTSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTSR 297

RESULT 7
 QYVBA PRELIMINARY; PRT; 298 AA.

AC QYVBA4;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Attachment glycoprotein G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-WV5222;
 RX MEDLINE=9022964; PubMed=9806017;
 RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.,
 RT "Antigenic and genetic diversity among the attachment proteins of
 RT group A respiratory syncytial viruses that have caused repeat
 RT infections in children."
 RL J. Infect. Dis. 178:925-932(1998).
 DR EMBL; AF065406; AAD02942.1; -
 DR InterPro; IPR000925; Glycoprot_G.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00802; Glycoprotein_G; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 SO SEQUENCE 298 AA; 32781 MW; AE7C9E05547C5745 CRC64;

Query Match 87.8%; Score 1356; DB 12; Length 298;
 Best Local Similarity 88.9%; Pred. No. 2.6e-88;
 Matches 265; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

QY 1 MSKNKQDRTAKTLEKMTDNLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60
 Db 1 MSKNKQDRTAKTLEKMTDNLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60
 QY 61 FIASANHKVLTLLTAIIODATSOIKNTTPYVLTQDPOLGISFNSLSEITTSQTTIIASTTP 120
 Db 61 FIASANHKVLTLLTAIIODATSOIKNTTPYVLTQNPOLGISFNSLSEITTSQTTIIASTTP 120
 QY 121 GYKSNLOPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEVEFNVPSCISGNNP 180
 Db 121 GYKSTLOSTTVGKTNTTTQAOQPNKPTTKORONKPPNKNDFEVEFNVPSCISGNNP 180

QY 181 TCMAICKRIIPNKKPGKKTTPKPKPTTKKDKLPQTTKPEVEPTTKPTEPTTINTTK 240
 Db 181 TCMAICKRIIPNKKPGKKTTPKPKPTTKKDKLPQTTKPEVEPTTKPTEPTTINTTK 240
 QY 241 TINTTLLTNNTGNPKLTSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTTR 298
 Db 241 TINTTLLTNTNRNPELTSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTTR 298

RESULT 8
 QYVBA PRELIMINARY; PRT; 297 AA.

AC Q82071;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein.
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SUBGROUP A;
 RX MEDLINE=94335057; PubMed=8057427;
 RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisile S., Russi J.,
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
 RA Melero J.A.;
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
 RT A): circulating lineages and correlation of genetic and antigenic
 RT changes in the G glycoprotein."
 RL J. Virol. 68:5448-5459(1994).
 DR EMBL; Z33432; CAA83875.1; -
 DR InterPro; IPR000925; Glycoprot_G.
 DR InterPro; IPR003880; Ppantne_attach.
 DR Pfam; PF00802; Glycoprotein_G; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
 SO SEQUENCE 297 AA; 32581 MW; AD947B2318193306 CRC64;

Query Match 87.7%; Score 1354; DB 12; Length 297;
 Best Local Similarity 89.2%; Pred. No. 3.6e-88;
 Matches 265; Conservative 10; Mismatches 22; Indels 0; Gaps 0;

QY 1 MSKNKQDRTAKTLEKMTDNLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60
 Db 1 MSKNKQDRTAKTLEKMTDNLNHLFISSGLYKLNKLSVAQITLSILAMITSLITAI 60
 QY 61 FIASANHKVLTLLTAIIODATSOIKNTTPYVLTQDPOLGISFNSLSEITTSQTTIIASTTP 120
 Db 61 FIASANHKVLTLLTAIIODATNOIKNTTPYVLTQNPOLGISFNSPSEITSLITIIIDSTTP 120
 QY 121 GYKSNLOPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEVEFNVPSCISGNNP 180
 Db 121 GYKSTLOSTTVGKTNTTTQAOQPNKPTTKORONKPPNKNDFEVEFNVPSCISGNNP 180
 QY 181 TCMAICKRIIPNKKPGKKTTPKPKPTTKKDKLPQTTKPEVEPTTKPTEPTTINTTK 240
 Db 181 TCMAICKRIIPNKKPGKKTTPKPKPTTKKDKLPQTTKPEVEPTTKPTEPTTINTTK 240
 QY 241 TINTTLLTNNTGNPKLTSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTTR 297
 Db 241 TINTTLLTNTNRNPELTSMETFHSTSEGNLSPSOVSTSEHPSQSPSPNTSR 297

RESULT 9
 QYVBA PRELIMINARY; PRT; 297 AA.
 AC Q82067;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE Glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUBGROUP A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrasile S., Russi J.,
RA Horita M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): correlating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:548-549(1994).
DR EMBL: Z33428; CA83871.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.
SQ SEQUENCE 297 AA; 32555 MW; 900A69BECAB66C CRC64;

Query Match 87.1%; Score 1345; DB 12; Length 297;
Best Local Similarity 88.9%; Pred. No. 1.6e-87;
Matches 264; Conservative 10; Mismatches 23; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKWTWTLNHLFISSGLYKLNKSAVQITSLIAMIISTSLITTAI 60
DB 1 MSKNDQRTAKTLEKWTWTLNHLFISSGLYKLNKSAVQITSLIAMIISTSLITTAI 60
OY 61 FIASANKVLTLLTALIODATSOIKNTPTLYLQDPOLGISFSNLSSETTSQTLLASTTP 120
DB 61 FIASANKVLTSTITLIDATNOIKNTPTLYLQDPOLGISFSNLSSETTSQTLLASTTP 120
OY 121 GVSNIQPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180
DB 121 GVSNIQPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180
OY 181 TCMAICKRIPKKRGGKTTTKPTKKPTTKKDKLPQTKPKREVPPTKPEEPTINTTK 240
DB 181 TCMAICKRIPKKRGGKTTTKPTKKPTTKKDKLPQTKPKREVPPTKPEEPTINTTK 240
OY 241 TMTTLLTNTGNGPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPNTT 297
DB 241 TMTTLLTNTGNGPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPNTT 297

RESULT 10
082058 PRELIMINARY; PRT; 297 AA.
AC 082058;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE (MAD-1-89) subgroup A, G glycoprotein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUBGROUP A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrasile S., Russi J.,
RA Horita M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): correlating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:548-549(1994).

DR EMBL: Z33456; CA83879.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.
SQ SEQUENCE 297 AA; 32741 MW; 0E567A174BE64964 CRC64;

Query Match 87.0%; Score 1344; DB 12; Length 297;
Best Local Similarity 88.5%; Pred. No. 1.8e-87;
Matches 262; Conservative 10; Mismatches 24; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKWTWTLNHLFISSGLYKLNKSAVQITSLIAMIISTSLITTAI 60
DB 1 MSKNDQRTAKTLEKWTWTLNHLFISSGLYKLNKSAVQITSLIAMIISTSLITTAI 60
OY 61 FIASANKVLTLLTALIODATSOIKNTPTLYLQDPOLGISFSNLSSETTSQTLLASTTP 120
DB 61 FIASANKVLTLLTALIODATSOIKNTPTLYLQDPOLGISFSNLSSETTSQTLLASTTP 120
OY 121 GVSNIQPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180
DB 121 SAESTPOSTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180
OY 181 TCMAICKRIPKKRGGKTTTKPTKKPTTKKDKLPQTKPKREVPPTKPEEPTINTTK 240
DB 181 TCMAICKRIPKKRGGKTTTKPTKKPTTKKDKLPQTKPKREVPPTKPEEPTINTTK 240
OY 241 TMTTLLTNTGNGPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPNTT 296
DB 241 TMTTLLTNTGNGPKLTSMETFHSTSESGNLSPOSVSTSEHPSQSSPNTT 296

RESULT 11
091947 PRELIMINARY; PRT; 297 AA.
AC 091947;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Attachment protein.
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CH34;
RX MEDLINE=98418504; PubMed=9747732;
RA Peret T.C.T., Hall C.B., Schenabel K.C., Golub J.A., Anderson L.J.;
RT "Circulation Patterns of Genetically Distinct Group A and B Strains of
RT Human Respiratory Syncytial Virus in a Community.";
RL J. Gen. Virol. 79:2221-2229(1998).
DR EMBL: AF065257; AAC36327.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.
SQ SEQUENCE 297 AA; 32564 MW; 377C7C8CDD0A3645 CRC64;

Query Match 86.7%; Score 1339; DB 12; Length 297;
Best Local Similarity 88.2%; Pred. No. 4.2e-87;
Matches 262; Conservative 11; Mismatches 24; Indels 0; Gaps 0;

OY 1 MSKNDQRTAKTLEKWTWTLNHLFISSGLYKLNKSAVQITSLIAMIISTSLITTAI 60
DB 1 MSKTKQGRKAKTLEKWTWTLNHLFISSGLYKLNKSAVQITSLIAMIISTSLITTAI 60
OY 61 FIASANKVLTLLTALIODATSOIKNTPTLYLQDPOLGISFSNLSSETTSQTLLASTTP 120
DB 61 FIASANKVLTSTITLIDATNOIKNTPTLYLQDPOLGISFSNLSSETTSQTLLASTTP 120
OY 121 GVSNIQPTVTKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNP 180

```

Db 121 GVKSTLSTVGTGKNTTQTAQPKPTTKORQNTPSKPNNDHFEVNFVPCISGNNP 180
Oy 181 TCMAICKRIPKPKGKTTTPTKPKPTTKDLKDPQTKPEVPTKPEPTINTTK 240
Db 181 TCMAICKRIPKPKGKTTTPTKPKPTTKDLKDPQTKPEVPTKPEPTINTTK 240
Oy 241 TMTTTLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPSPPTTR 297
Db 241 TMTTTLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPSPPTTR 297

RESULT 12
OyYVB3 PRELIMINARY: PRT: 293 AA.
ID 09YVB3
AC 09YVB3;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)
DE Attachment glycoprotein G (Fragment).
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WB6973;
RX MEDLINE=99022964; PubMed=9806017;
RA Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;
RT "Antigenic and genetic diversity among the attachment proteins of
RT group A respiratory syncytial viruses that have caused repeat
RT infections in children.";
RL J. Infect. Dis. 178:925-932(1998).
DR EMBL; AF065407; AAD02943.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR InterPro: IPR002965; P_rich_extensn.
DR Pfam: PF00802; Glycoprotein_G. 1.
DR PRINTS; PR01217; PRICHEXTENS.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 293 AA; 32125 MW; 17B5B43396A63CFC CRC64;

Query Match 86.5%; Score 1336; DB 12; Length 293;
Best Local Similarity 90.1%; Pred. No. 6.7e-87;
Matches 264; Conservative 7; Mismatches 22; Indels 0; Gaps 0;

Oy 6 DQRTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSIAMIISTSLITAIIFASA 65
Db 1 DQRTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSIAMIISTSLITAIIFASA 60
Oy 66 NKKVLTTLTIODATSOIKNTPTYLTOPOLGISFSNLSSETTSOTTLASTPVGKSN 125
Db 61 NKKVLTTLTIODATSOIKNTPTYLTOPOLGISFSNLSSETTSOTTLASTPVGKST 120
Oy 126 LQPTTKTKNTTTOQPSKPTTKORQNPKNPNNDHFEVNFVPCISGNNPTCAI 185
Db 121 LQPTTKTKNTTTOQPSKPTTKORQNPKNPNNDHFEVNFVPCISGNNPTCAI 180
Oy 186 CKRIPKPKGKTTTPTKPKPTTKDLKDPQTKPEVPTKPEPTINTTKNTT 245
Db 181 CKRIPKPKGKTTTPTKPKPTTKDLKDPQTKPEVPTKPEPTINTTKNTT 240
Oy 246 TLTNTTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPSPPTTRQ 298
Db 241 TLTNTTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPSPPTTRQ 293

RESULT 13
OyYVB3 PRELIMINARY: PRT: 295 AA.
ID 086359
AC 086359;
DT 01-NOV-1996 (TReMBLrel. 01, Created)

```

```

DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE G protein (Fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-6256;
RA Cane P.A.;
RX Submitted (Apr-1993) to the EMBL/Genbank/DBJ databases.
RL EMBL: X73353; CAA51764.1;
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003880; Pantine_attach.
DR Pfam: PF00802; Glycoprotein_G. 1.
DR PROSITE; PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match 86.4%; Score 1344; DB 12; Length 295;
Best Local Similarity 88.4%; Pred. No. 3.9e-87;
Matches 260; Conservative 8; Mismatches 26; Indels 0; Gaps 0;

Oy 3 KNDQRTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSIAMIISTSLITAIIFI 62
Db 1 KNDQRTAKTEKTWDTLNHLFISSGLYKLNKSAQITLSIAMIISTSLITAIIFI 60
Oy 63 ASANHKVLTTLTIODATSOIKNTPTYLTOPOLGISFSNLSSETTSOTTLASTPVG 122
Db 61 ASANHKVLTTLTIODATSOIKNTPTYLTOPOLGISFSNLSSETTSOTTLASTPVG 120
Oy 123 KSNLQPTTKTKNTTTOQPSKPTTKORQNPKNPNNDHFEVNFVPCISGNNPTC 182
Db 121 ESTPQSTTVTKNTTTOQPSKPTTKORQNPKNPNNDHFEVNFVPCISGNNPTC 180
Oy 183 WAIKCRIPKPKGKTTTPTKPKPTTKDLKDPQTKPEVPTKPEPTINTTKTN 242
Db 181 WAIKCRIPKPKGKTTTPTKPKPTTKDLKDPQTKPEVPTKPEPTINTTKTN 240
Oy 243 ITTTLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPSPPTTR 296
Db 241 ITTTLTNNNTGNPKLTSOMETFSHSTSSGNNLSPSOVSTSHSPSPPTTR 294

RESULT 14
OyYVB3 PRELIMINARY: PRT: 295 AA.
ID 086360
AC 086360;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, last annotation update)
DE G protein (Fragment).
OS Respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=12814;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RSB89-642;
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).

```

[2]
 RN SEQUENCE FROM N.A.
 RP STRAIN-RSB89-642;
 RA Cane P.A.;
 RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
 DR EMBL: XJ3334; CAAS1765.1; -;
 DR InterPro: IPR000925; Glycoprot_G.
 DR InterPro: IPR003880; Pplantn.attach.
 DR Pfam: PF00802; Glycoprotein.G; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 295 AA; 32306 MW; 671DCF041570284 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 295;
 Best Local Similarity 88.5%; Pred. No. 9,3e-87;
 Matches 261; Conservative 11; Mismatches 23; Indels 0; Gaps 0;

QY 3 KNDQRTAKTEKTWDTLNLHLFISSGLYKLNKSAQITLSILAMISTSLITAIIFI 62
 DB 1 KNDQRTAKTEKTWDTLNLHLFISSGLYKLNKSAQITLSILAMISTSLITAIIFI 60
 QY 63 ASANHKVTLITATIODATSOIKNTPTVLTODPOLGISFNSLSEITTSOTTTILASTTPGV 122
 DB 61 ASANHKITSTTTIIOATNOIKNTPTVLTQNPOLGISFNSNSDITSLITLIDSTTPGV 120
 QY 123 KSNLQPTTVKTKNTTTOTOPSKPTTKORONKPPNNDPHEVEFVPCISCSNPTC 182
 DB 121 KSTLQSTVGTKNVTTTQOPNKPPTTKORONKPPNNDPHEVEFVPCISCSNPTC 180
 QY 183 MAICKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPEVPTTKPEEPTINTTKTN 242
 DB 181 MAICKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPEVPTTKPEEPTINTTKTN 240
 QY 243 ITTTLTNTNNTGNPKLTSOMETFEHSTSGNLSPOSVSTSEHPSOPSSPMTTR 297
 DB 241 ITTTLTNTNNTGNPKLTSOMETFEHSTSGNLSPOSVSTSEHPSOPSSPMTTR 295

RESULT 15

ID 082065 PRELIMINARY; PRT: 298 AA.
 AC 082065;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein.
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SUBGROUP A;
 RX MEDLINE=94335057; PubMed=8057427;
 RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrasile S., Russi J.,
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
 RA Melero J.A.;
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
 RT A): cocirculating lineages and correlation of genetic and antigenic
 RT changes in the G glycoprotein.";
 RL J. Virol. 68:5448-5459(1994).
 DR EMBL: Z33426; CAA83869.1; -;
 DR InterPro: IPR000925; Glycoprot_G.
 DR InterPro: IPR003880; Pplantn.attach.
 DR Pfam: PF00802; Glycoprotein.G; 1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 SQ SEQUENCE 298 AA; 32729 MW; 4E90FPD43B45744 CRC64;

Query Match 86.4%; Score 1334; DB 12; Length 298;
 Best Local Similarity 88.2%; Pred. No. 9,4e-87;
 Matches 261; Conservative 10; Mismatches 25; Indels 0; Gaps 0;

QY 1 MSKNQDRTAKTEKTWDTLNLHLFISSGLYKLNKSAQITLSILAMISTSLITAIIFI 60
 DB 1 MSKNQDRTAKTEKTWDTLNLHLFISSGLYKLNKSAQITLSILAMISTSLITAIIFI 60
 QY 61 FIASANHKVTLITATIODATSOIKNTPTVLTODPOLGISFNSLSEITTSOTTTILASTTP 120
 DB 61 FIASANHKVTLITATIODATSOIKNTPTVLTQNPOLGISFNSLSEITTSOTTTILASTTP 120
 QY 121 GVSMLQPTTVKTKNTTTOTOPSKPTTKORONKPPNNDPHEVEFVPCISCSNPTC 180
 DB 121 SAESTPLSTVTKNTTTTQOPSKPTTKORONKPPNNDPHEVEFVPCISCSNPTC 180
 QY 181 TCMAICKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPEVPTTKPEEPTINTTKTN 240
 DB 181 TCMAICKRIPNKKRGKKTTPKPKPTTKKDLKPOTTKPEVPTTKPEEPTINTTKTN 240
 QY 241 TMTTLTNTNNTGNPKLTSOMETFEHSTSGNLSPOSVSTSEHPSOPSSPMTTR 296
 DB 241 TMTTLTNTNNTGNPKLTSOMETFEHSTSGNLSPOSVSTSEHPSOPSSPMTTR 296

Search completed: May 1, 2003, 13:46:23
 Job time : 36.9849 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:55:01 ; Search time 1442.68 Seconds

(without alignments)
14423.488 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715
Sequence: 1 cacaagtcacactaacac.....gtacgtatttaaaaaaaaaa 715

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rpd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	920	6 A16257	A16257 Synthetic R
2	715	100.0	920	6 A16258	A16258 Synthetic R
3	715	100.0	920	6 AR080406	AR080406 Sequence
4	715	100.0	920	6 AR092530	AR092530 Sequence
5	715	100.0	920	6 AR122885	AR122885 Sequence
6	715	100.0	920	6 AR123857	AR123857 Sequence
7	715	100.0	920	6 AR148357	AR148357 Sequence
8	710	99.3	923	6 HRSRMAG	X17085 Human resp1
9	709	99.2	922	14 HRSVGL16	Z33429 Human resp1
10	699.2	97.8	917	14 RSHGLYG	M17212 Human resp1
11	696	97.3	894	6 AR080424	AR080424 Sequence
12	696	97.3	894	6 AR092548	AR092548 Sequence
13	654.8	91.6	8510	14 RSHICE	M11466 Human resp1
14	654.8	91.6	15222	6 AR093219	AR093219 Sequence
15	654.8	91.6	15222	14 HRU50362	U50362 Human resp1
16	654.8	91.6	15222	14 HRU50363	U50363 Human resp1
17	654.8	91.6	15222	14 HRU63644	U63644 Human resp1
18	654.8	91.6	15222	14 RSHSPD	M74568 Human resp1
19	654.8	91.6	15223	6 AR089137	AR089137 Sequence
20	654.8	91.6	15223	14 AF035006	AF035006 Human res
21	649.6	90.9	914	14 PARSENG	X03149 Respiratory
22	649.6	90.9	918	6 AX047075	AX047075 Sequence
23	646.2	90.4	897	6 AX339011	AX339011 Sequence
24	641.8	89.8	15190	14 HRU39662	X39662 Human resp1
25	641.8	89.8	15191	14 RSU39661	U39661 Respiratory
26	628.6	87.9	840	14 AF193308	AF193308 Human res
27	628.2	87.9	918	14 AF065406	AF065406 Human res
28	625.4	87.5	840	14 AF193306	AF193306 Human res
29	621	86.9	922	14 HRSVCV3	Z33412 Human resp1
30	621	86.9	922	14 HRSVCV6	Z33456 Human resp1
31	621	86.9	922	14 HRSVGL11	X33474 Human resp1
32	621	86.9	922	14 HRSVGL12	X33475 Human resp1
33	620.8	86.8	887	14 AF065409	AF065409 Human res
34	620.2	86.7	918	14 AF065405	AF065405 Human res
35	619.4	86.6	922	14 HRSVGL2	Z33422 Human resp1
36	618.4	86.5	922	14 HRSVGL1	Z33410 Human resp1
37	618.4	86.5	922	14 HRSVGL12	Z33454 Human resp1
38	618.4	86.5	922	14 HRSVGL12	Z33494 Human resp1
39	617.8	86.4	922	14 AF065256	AF065256 Human res
40	617.8	86.4	922	14 HRSVGL13	AF065256 Human res
41	617.8	86.4	922	14 HRSVGL9	Z33426 Human resp1
42	616.8	86.4	922	14 HRSVGL9	Z33422 Human resp1
43	616.8	86.3	922	14 HRSVGLX1	Z33493 Human resp1
44	616.2	86.2	922	14 HRSVGL3	Z33416 Human resp1
45	616.2	86.2	922	14 HRSVGL4	Z33417 Human resp1

ALIGNMENTS

RESULT 1
A16257 LOCUS A16257 920 bp DNA
DEFINITION Synthetic RSV G gene (seq ID No: 7).
ACCESSION A16257
VERSION A16257.1 GI:640933
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
FEATURES
source Location/Qualifiers
1..920
/organism="synthetic construct"
/db_xref="taxon:32650"
8..904
gene

Pred. No. is the number of results predicted by chance to have a

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL Patent: US 5968776-A 7 19-OCT-1999;
FEATURES
source location/Qualifiers
1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 715; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.6e-149;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAAGTCACACTAACAAGTCAATCATACAGATGCACAGCCAGATCAAGACACA 60
DB 206 CACAAAGTCACACTAACAAGTCAATCATACAGATGCACAGCCAGATCAAGACACA 265
QY 61 ACCCAACATACCTCTACTAGAGATCTCTAGCTTGGAATCAGCTTCTCCAACTGTCTGAA 120
DB 266 ACCCAACATACCTCTACTAGAGATCTCTAGCTTGGAATCAGCTTCTCCAACTGTCTGAA 325
QY 121 ATTACATCAACAAGCCACCATCTACTAGCTTCAACACACAGAGATCAAGTCAAACTG 180
DB 326 ATTACATCAACAAGCCACCATCTACTAGCTTCAACACACAGAGATCAAGTCAAACTG 385
QY 181 CAACCCACACAGTCAAGACTAATAAACAACAACAACCAACCAACCAACCAACCAACCC 240
DB 386 CAACCCACACAGTCAAGACTAATAAACAACAACAACCAACCAACCAACCAACCAACCC 445
QY 241 ACTACAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGAGCAACAATCAACCTGTGGCTATCTGC 360
DB 506 GTGTTTAACTTTGTACCTGAGCATATGAGAGCAACAATCAACCTGTGGCTATCTGC 565
QY 361 AAAAGAAATACCAAAACAAAAACAGAAAGAAAAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGAAATACCAAAACAAAAACAGAAAGAAAAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAAGACAAACAAAAAGATCTCAAACTCTAAACCACTAAACCAAGAAAGTACCC 480
DB 626 ACCTTCAAGACAAACAAAAAGATCTCAAACTCTAAACCACTAAACCAAGAAAGTACCC 685
QY 481 ACCCAAGCCCAAGAGAGAGCCCAACCATCAAGACCAACCAACCAACCAACCAACCA 540
DB 686 ACCCAAGCCCAAGAGAGAGCCCAACCATCAAGACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTACCAACAACACACAGAGAAATCCAAAATCAAGTCAAAATGGAATCTTCAC 600
DB 746 CTGCTACCAACAACACACAGAGAAATCCAAAATCAAGTCAAAATGGAATCTTCAC 805
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCACAAACATCCGACACCA 660
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCACAAACATCCGACACCA 865
QY 661 TCACAACCTCTATCTCACCACCAACACAGCCAGTAGTATTATTAATAAAAAAAA 715
DB 866 TCACAACCTCTATCTCACCACCAACACAGCCAGTAGTATTATTAATAAAAAAAA 920
RESULT 4
LOCUS AR092530 920 bp DNA linear PAT 08-SEP-2000
DEFINITION Sequence 7 from patent US 599169.
ACCESSION AR092530

VERSION AR092530.1 GI:10019284
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasyszyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL Patent: US 5998169-A 7 07-DEC-1999;
FEATURES
source location/Qualifiers
1..920
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN
Query Match 100.0%; Score 715; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.6e-149;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAAGTCACACTAACAAGTCAATCATACAGATGCACAGCCAGATCAAGACACA 60
DB 206 CACAAAGTCACACTAACAAGTCAATCATACAGATGCACAGCCAGATCAAGACACA 265
QY 61 ACCCAACATACCTCTACTAGAGATCTCTAGCTTGGAATCAGCTTCTCCAACTGTCTGAA 120
DB 266 ACCCAACATACCTCTACTAGAGATCTCTAGCTTGGAATCAGCTTCTCCAACTGTCTGAA 325
QY 121 ATTACATCAACAAGCCACCATCTACTAGCTTCAACACACAGAGATCAAGTCAAACTG 180
DB 326 ATTACATCAACAAGCCACCATCTACTAGCTTCAACACACAGAGATCAAGTCAAACTG 385
QY 181 CAACCCACACAGTCAAGACTAATAAACAACAACAACCAACCAACCAACCAACCAACCC 240
DB 386 CAACCCACACAGTCAAGACTAATAAACAACAACAACCAACCAACCAACCAACCAACCC 445
QY 241 ACTACAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAACACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTTAACTTTGTACCTGAGCATATGAGAGCAACAATCAACCTGTGCTATCTGC 360
DB 506 GTGTTTAACTTTGTACCTGAGCATATGAGAGCAACAATCAACCTGTGCTATCTGC 565
QY 361 AAAAGAAATACCAAAACAAAAACAGAAAGAAAAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGAAATACCAAAACAAAAACAGAAAGAAAAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAAGACAAACAAAAAGATCTCAAACTCTAAACCACTAAACCAAGAAAGTACCC 480
DB 626 ACCTTCAAGACAAACAAAAAGATCTCAAACTCTAAACCACTAAACCAAGAAAGTACCC 685
QY 481 ACCCAAGCCCAAGAGAGCCCAACCATCAAGACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACCCAAGCCCAAGAGAGCCCAACCATCAAGACCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTACCAACAACACACAGAGAAATCCAAAATCAAGTCAAAATGGAATCTTCAC 600
DB 746 CTGCTACCAACAACACACAGAGAAATCCAAAATCAAGTCAAAATGGAATCTTCAC 805
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCACAAACATCCGACACCA 660
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTTCACAAACATCCGACACCA 865
QY 661 TCACAACCTCTATCTCACCACCAACACAGCCAGTAGTATTATTAATAAAAAAAA 715
DB 866 TCACAACCTCTATCTCACCACCAACACAGCCAGTAGTATTATTAATAAAAAAAA 920
RESULT 5
LOCUS AR122885 920 bp DNA linear PAT 16-MAY-2001
DEFINITION
ACCESSION AR122885

DEFINITION Sequence 7 from patent US 6168786.
ACCESSION ARI22885
VERSION ARI22885.1 GI:14107851
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 920)
AUTHORS Klein, M.H., Du, R.-P. and Ewasyszyn, M.E.
TITLE Chimeric immunogens
JOURNAL Patent: US 6168786 A 7 02-JAN-2001;
FEATURES
source Location/Qualifiers
1..920
/organism="unknown"
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN

Query Match 100.0%; Score 715; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.6e-149;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAATCACTGCAATCATACAGATGCAACAGCGAGATCAAGAACACA 60
DB 206 CACAAGTCACACTAATCACTGCAATCATACAGATGCAACAGCGAGATCAAGAACACA 265
QY 61 ACCCAACATACCTCACTCAGATCTCAGCTTGGAAATCAGCTTCTCAATCTGTCTGAA 120
DB 266 ACCCAACATACCTCACTCAGATCTCAGCTTGGAAATCAGCTTCTCAATCTGTCTGAA 325
QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACAAACCAAGAGTCAAGTCAAACTG 180
DB 326 ATTACATCAACAAACCAACCACTACTAGCTTCAACAAACCAAGAGTCAAGTCAAACTG 385
QY 181 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAACCAACCC 240
DB 386 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAACCC 445
QY 241 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 300
DB 446 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 505
QY 301 GTGTTTAACCTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 360
DB 506 GTGTTTAACCTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 565
QY 361 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAACCAAGCTCAAAAAAACCA 420
DB 566 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAACCAAGCTCAAAAAAACCA 625
QY 421 ACCTTCAAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGTACCC 480
DB 626 ACCTTCAAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGTACCC 685
QY 481 ACCACCAAGCCCAAGAAAGCCAAACATCAACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCCAAGAAAGCCAAACATCAACCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTCACCACAAACACAGAGAAATCCAAACTCAAAAGTCAAAATGAAACCTTCCAC 600
DB 746 CTGCTCACCACAAACACAGAGAAATCCAAACTCAAAAGTCAAAATGAAACCTTCCAC 805
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGACCCA 660
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGACCCA 865
QY 661 TCACAAACCTCATCTCCACCCACACACACAGCCAGTAGTATTATTAACAAAAA 715
DB 866 TCACAAACCTCATCTCCACCCACACACACAGCCAGTAGTATTATTAACAAAAA 920

RESULT 6
ARI23540
LOCUS ARI23540 920 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6171783.
ACCESSION ARI23540
VERSION ARI23540.1 GI:14108901
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 920)
AUTHORS Klein, M.H., Du, R.-P. and Ewasyszyn, M.E.
TITLE Infection detection method using chimeric protein
JOURNAL Patent: US 6171783-A 7 09-JAN-2001;
FEATURES
source Location/Qualifiers
1..920
/organism="unknown"
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN

Query Match 100.0%; Score 715; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.6e-149;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAGTCACACTAATCACTGCAATCATACAGATGCAACAGCGAGATCAAGAACACA 60
DB 206 CACAAGTCACACTAATCACTGCAATCATACAGATGCAACAGCGAGATCAAGAACACA 265
QY 61 ACCCAACATACCTCACTCAGATCTCAGCTTGGAAATCAGCTTCTCAATCTGTCTGAA 120
DB 266 ACCCAACATACCTCACTCAGATCTCAGCTTGGAAATCAGCTTCTCAATCTGTCTGAA 325
QY 121 ATTACATCAACAAACCAACCACTACTAGCTTCAACAAACCAAGAGTCAAGTCAAACTG 180
DB 326 ATTACATCAACAAACCAACCACTACTAGCTTCAACAAACCAAGAGTCAAGTCAAACTG 385
QY 181 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAACCC 240
DB 386 CAACCCACACAGTCAAGACTAAAAACACACACACCAACCAACCAACCAACCAACCC 445
QY 241 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 300
DB 446 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 505
QY 301 GTGTTTAACCTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 360
DB 506 GTGTTTAACCTTGTACCTGCGACATATGACAGACAAATCCAACTGCTGGCTATCTGC 565
QY 361 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAACCAAGCTCAAAAAAACCA 420
DB 566 AAAAGAATACCAAAACCAAGAGAAAGAAACCAACCAACCAAGCTCAAAAAAACCA 625
QY 421 ACCTTCAAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGTACCC 480
DB 626 ACCTTCAAGACAAACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAGTACCC 685
QY 481 ACCACCAAGCCCAAGAAAGCCAAACATCAACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCCAAGAAAGCCAAACATCAACCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTCACCACAAACACAGAGAAATCCAAACTCAAAAGTCAAAATGAAACCTTCCAC 600
DB 746 CTGCTCACCACAAACACAGAGAAATCCAAACTCAAAAGTCAAAATGAAACCTTCCAC 805
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGACCCA 660
DB 806 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCGAGACCCA 865
QY 661 TCACAAACCTCATCTCCACCCACACACACAGCCAGTAGTATTATTAACAAAAA 715
DB 866 TCACAAACCTCATCTCCACCCACACACACAGCCAGTAGTATTATTAACAAAAA 920

RESULT 7
ARI48357
LOCUS ARI48357 920 bp DNA linear PAT 08-AUG-2001

DEFINITION Sequence 7 from patent US 6225091.
LOCUS ARI48357
VERSION ARI48357.1 GI:15112447
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 920)
AUTHORS Klein,M.H., Du,R.-P. and Ewasysbyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL Patent: US 6225091-A 7 01-MAY-2001;
FEATURES
source Location/Qualifiers
BASE COUNT 380 a 290 c 95 g 155 t
ORIGIN

Query Match 100.0%; Score 715; DB 6; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.6e-149;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTAACAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60
DB 206 CACAAAGTACACTAACAACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 265
QY 61 ACCCCACATACCTCAGCTAGAGTCTGAGTGGAAATGAGTCTGCCAATGCTGTA 120
DB 266 ACCCCACATACCTCAGCTAGAGTCTGAGTGGAAATGAGTCTGCCAATGCTGTA 325
QY 121 ATTACATGCAAAACGACACCATAGTCTGACAGACAGGAGTCAAGTCAAACTG 180
DB 326 ATTACATGCAAAACGACACCATAGTCTGACAGACAGGAGTCAAGTCAAACTG 385
QY 181 CAACCCACAGAGTCAAGACTAAAAACAGACACACCAACCCAAACCAACGACGACC 240
DB 386 CAACCCACAGAGTCAAGACTAAAAACAGACACACCAACCCAAACCAACGACGACC 445
QY 241 ACTACAAAACAGGCAAAACCAACACCAACCAACCAATATGATTTTCATTGGAA 300
DB 446 ACTACAAAACAGGCAAAACCAACACCAACCAACCAATATGATTTTCATTGGAA 505
QY 301 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 360
DB 506 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 565
QY 361 AAAAGAAATACCAAAACCAAGGAAAGAAACACACGCTACGACAAAAACCA 420
DB 566 AAAAGAAATACCAAAACCAAGGAAAGAAACACACGCTACGACAAAAACCA 625
QY 421 ACCTTCAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAAAGGAAATACCC 480
DB 626 ACCTTCAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAAAGGAAATACCC 685
QY 481 ACCACCAAGCCACAGAGAGCCACATCAACACACCAAAACCAACATCAACAATCA 540
DB 686 ACCACCAAGCCACAGAGAGCCACATCAACACACCAAAACCAACATCAACAATCA 745
QY 541 CTGCTCACCACACACACAGAGAAATGCAAACTCAAGTCAATGGAATGGAACCTTCAC 600
DB 746 CTGCTCACCACACACACAGAGAAATGCAAACTCAAGTCAATGGAATGGAACCTTCAC 805
QY 601 TCAACCTCTCCGAGGCAATCTAAGCCTTCTCAAGTCTCCACAACTCGAGACACCA 660
DB 806 TCAACCTCTCCGAGGCAATCTAAGCCTTCTCAAGTCTCCACAACTCGAGACACCA 865
QY 661 TCACAACCTCATCTCAACCAACCAACAGCCAGTAGTATTAAAAAAGAAAAA 715
DB 866 TCACAACCTCATCTCAACCAACCAACAGCCAGTAGTATTAAAAAAGAAAAA 920

HRSRNAG
LOCUS HSRNAG 923 bp mRNA linear VRL 24-JUN-1992
DEFINITION Human respiratory syncytial (RS) virus mRNA for G protein.
ACCESSION X17085
VERSION X17085.1 GI:60306
KEYWORDS g protein.
SOURCE Human respiratory syncytial virus.
ORGANISM Human respiratory syncytial virus.
REFERENCE Viruses: ssRNA negative-strand viruses: Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.
1 (bases 1 to 923)
AUTHORS Garcia-Barreno,B., Portela,A., Delgado,T., Lopez,J.A. and Melero,J.A.
TITLE Frame shift mutations as a novel mechanism for the generation of neutralization resistant mutants of human respiratory syncytial virus
JOURNAL EMBO J. 9 (12), 4181-4187 (1990)
MEDLINE 91065351
PUBMED 2249671
FEATURES
source Location/Qualifiers
1..923
/organism="Human respiratory syncytial virus";
/db_xref="taxon:11250"
16..912
/codon_start=1
/product="G protein"
/protein_id="CAA34937.1"
/db_xref="GI:60307"
/translation="MSKNDQRTAKLTLEKTDWTLNLYLFISGGLYNLKSIAQITLS ILMIISTSLIITAIIFIASANRVTLLTAIIODATSOIKNTTPTVLDDPOLGISFS NLSITSTSTLIATSTPGVSKSLQLOPTKNTTTOPTSPPTKORONPKPN NDHPFVFNVPVCSISNNPTCAICRIKIPNPKGKKTTRKPTKPTFKTDLKPO TPKPKVPTTKPEPTPTKNTKNTITTTLLTNNTGNPLITQSMETFHSTSECNLSP SSVSTSEHSDSPSPNTTTO"
BASE COUNT 379 a 290 c 98 g 156 t
ORIGIN

Query Match 99.3%; Score 710; DB 14; Length 923;
Best Local Similarity 100.0%; Pred. No. 2e-148;
Matches 710; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60
DB 214 CACAAAGTACACTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGACACA 273
QY 61 ACCCCACATACCTCAGCTAGAGTCTGACGCTTGAATGACGTTCTCCAAATCTGTGAA 120
DB 274 ACCCCACATACCTCAGCTAGAGTCTGACGCTTGAATGACGTTCTCCAAATCTGTGAA 333
QY 121 ATTACATGCAAAACGACACCATAGTCTGACAGACACGAGAGTCAAGTCAAACTG 180
DB 334 ATTACATGCAAAACGACACCATAGTCTGACAGACACGAGAGTCAAGTCAAACTG 393
QY 181 CAACCCACAGAGTCAAGACTAAAAACCAACCAACCAACCAACCAACCAACCAAGCC 240
DB 394 CAACCCACAGAGTCAAGACTAAAAACCAACCAACCAACCAACCAACCAACCAAGCC 453
QY 241 ACTACAAAACAGGCAAAACCAACCAACCAACCAACCAATATGATTTTCATTGGAA 300
DB 454 ACTACAAAACAGGCAAAACCAACCAACCAACCAACCAATATGATTTTCATTGGAA 513
QY 301 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 360
DB 514 GTGTTAACTTGTACCTGACGATATGACAGCAACATCCAACTGCTGGCTATCTGC 573
QY 361 AAAAGAAATACCAAAACCAAGGAAAGAAACCAACCAACCAAGCTTACAAAAAACA 420
DB 574 AAAAGAAATACCAAAACCAAGGAAAGAAACCAACCAACCAAGCTTACAAAAAACA 633
QY 421 ACCTTCAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAAAGGAAATACCC 480
DB 634 ACCTTCAGACACCAAAAGATCTCAAACTCAAACTCAAACTCAAAAGGAAATACCC 693

QY 481 ACCACCAAGCCACAGAGGCAACCATCAACACCAACCAACATCAACTACA 540
DB 694 ACCACCAAGCCACAGAGGCAACCATCAACACCAACCAACATCAACTACA 753
QY 541 CTGCTCACCACAAACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 600
DB 754 CTGCTCACCACAAACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 813
QY 601 TCACACTCTCCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCAGACCA 660
DB 814 TCACACTCTCCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCAGACCA 873
QY 661 TCACAACTCATCTCCACCAACACAGAGCAAGTCTTAATTAATAA 710
DB 874 TCACAACTCATCTCCACCAACACAGAGCAAGTCTTAATTAATAA 923

RESULT 9
HRSVGL16 922 bp RNA linear VRL 05-JUN-1997
LOCUS Human respiratory syncytial virus (MON-7-91) subgroup A, G gene for glycoprotein.
DEFINITION
ACCESSION 233429
VERSION 233429.1 GI:485888
KEYWORDS G gene; glycoprotein.
SOURCE Human respiratory syncytial virus.
ORGANISM Human respiratory syncytial virus.
REFERENCES: ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.
1 (bases 1 to 922)
Dopazo, J.
AUTHORS Direct Submission
TITLE Submitted (10-MAY-1994) Dopazo, J., Centro Nacional de Biotecnologia
JOURNAL CSIC, Bioinformatica, Universidad Autonoma, Cantoblanco, Madrid, SPAIN, 28049
2 (bases 1 to 922)
Garcia, O., Martin, M., Dopazo, J., Arbizu, J., Fabras, S., Russi, J., Hortel, M., Perez-Brena, P., Martinez, I., Garcia-Barreno, B. and Melero, J. A.
TITLE Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein
JOURNAL J. Virol. 68 (9), 5448-5459 (1994)
MEDLINE 9433057
PUBMED 8057427

FEATURES
Source
Location/Qualifiers
1..922
/organism="Human respiratory syncytial virus"
/strain="subgroup A"
/isolate="MON 7 91 (Montevideo/Uruguay, 1991)"
/db_xref="taxon:11250"
/cell_line="HeP-2"
16..912
/gene="G"
16..912
/partial
/gene="G"
/codon_start=1
/product="glycoprotein"
/protein_id="CAA83872.1"
/db_xref="GI:485888"
/db_xref="SPTREMBL:O82068"
/translation="ISKNDQPTAKTLEKWTDLNLYLFISSGLYLNKISIAQITLS ILAMISTLITAIIFIASANHKVLTATLADTSOIKNTPTLYLRODOLGISFS NLSITSCPTTILASTPTGYSNLOPTVKTKNTTGTQOSKPTTKROROKPKPKPD NDFHEVFNFVSCISGNNPTMAICKIPKPKPKTKTTTPTKPTTKTKDLKPD TTKPEVPTPTKPEPTINTKTNTITLLTNNNTGPKLTSQMETFSTSEGLSP SOVSTSEHPSPPNTRQ"
BASE COUNT 378 a 291 c 97 g 156 t
ORIGIN
Query Match 99.2%; Score 709; DB 14; Length 922;

Best Local Similarity 100.0%; Pred. No. 3,4e-148;
Matches 709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAAGTCACACTAACACTGCAATCATACATCAACCAAGCCGATCAAGAACCA 60
DB 214 CACAAAGTCACACTAACACTGCAATCATACATCAACCAAGCCGATCAAGAACCA 273
QY 61 ACCCAACATACCTACTCAGATGCTCAGCTTGAATTCAGCTTCTGCTGAA 120
DB 274 ACCCAACATACCTACTCAGATGCTCAGCTTGAATTCAGCTTCTGCTGAA 333
QY 121 ATTACATCAACACACACCATCTAGCTTGAATTCAGCTTCTGCTGAA 180
DB 334 ATTACATCAACACACACCATCTAGCTTGAATTCAGCTTCTGCTGAA 393
QY 181 CAACCCAAACAGTCAGACTTAAACACACCAACCAACCAACCAACCAACCAACCC 240
DB 394 CAACCCAAACAGTCAGACTTAAACACACCAACCAACCAACCAACCAACCAACCC 453
QY 241 ACTACAAACAGTCAGACTTAAACACCAACCAACCAACCAACCAACCAACCAACCC 300
DB 454 ACTACAAACAGTCAGACTTAAACACCAACCAACCAACCAACCAACCAACCAACCC 513
QY 301 GTGTTTAACTTTGTAACCTCGAGATATGCAACAATCCACCTGCTGCTATGCTG 360
DB 514 GTGTTTAACTTTGTAACCTCGAGATATGCAACAATCCACCTGCTGCTATGCTG 573
QY 361 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 574 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 634 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693
QY 481 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCA 540
DB 694 ACCACCAAGCCACAGAGAGCCCAACCATCAACCAACCAACCAACCAACCAACCA 753
QY 541 CTGCTCACCACAAACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 600
DB 754 CTGCTCACCACAAACACAGAGAAATCCAAAAGTCCAAATGGAACCTTCCAC 813
QY 601 TCACACTCTCCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCAGACCA 660
DB 814 TCACACTCTCCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCAAACATCCAGACCA 873
QY 661 TCACAACTCATCTCCACCAACACAGAGCAAGTCTTAATTAATAA 709
DB 874 TCACAACTCATCTCCACCAACACAGAGCAAGTCTTAATTAATAA 922

RESULT 10
RSHGLYG 917 bp ss-RNA linear VRL 03-AUG-1993
LOCUS Human respiratory syncytial virus (subgroup A) attachment protein
DEFINITION (g) mRNA, complete cds.
ACCESSION M17212
VERSION M17212.1 GI:333940
KEYWORDS attachment glycoprotein; surface glycoprotein.
SOURCE Human respiratory syncytial virus.
ORGANISM Human respiratory syncytial virus.
REFERENCES: ssRNA negative-strand viruses; Mononegavirales; Paramyxoviridae; Pneumovirinae; Pneumovirus.
1 (bases 1 to 917)
Johnson, P. R., Springs, M. K., Olmsted, R. A. and Collins, P. L.
AUTHORS The G glycoprotein of human respiratory syncytial viruses of subgroups A and B: extensive sequence divergence between antigenically related proteins
PROC. NATL. ACAD. SCI. U.S.A. 84 (16), 5625-5629 (1987)
JOURNAL MEDLINE 87289657
PUBMED 2441388

COMMENT The exact 5' end of long G mRNA was not determined.
FEATURES
Source location/Qualifiers
1..917
/organism="Human respiratory syncytial virus"
/db_xref="taxon:11250"
16..912
/note="attachment glycoprotein (G)"
/codon_start=1
/protein_id="AAA47411.1"
/db_xref="GI:333941"
/translation="MSKNDORTAKTLEKTMPTLHLLFISGLYKLNLSIAQTLLS
ILAMTISLITAIPIFASNNKKTLLTALIODTSQIKNTPTLYLTDPLGISFS
NLSEITSTQTTILASTPQVKSNIQPTTKNTKTNTTQIQSPKPTTKQKRNKPKPN
NDHEFENFVPCISISNNPTCMAICKRIPNKPKGKFTTKTKPKTKTKDKHPQ
TTKPKFVPTTKPEPTLNTTKNTLITLLTNNTGNPKLTSOMETPSTSEGNLSP
SOVSTSEHPSPSPPTTRQ"
BASE COUNT 375 a 290 c 97 g 154 t 1 others
ORIGIN
Query Match 97.8%; Score 699.2; DB 14; Length 917;
Best Local Similarity 99.6%; Pred. No. 5,2e-146;
Matches 701; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CACAAGTACACTACACTGCAATCATGCAAGATGCAACAAGCCAGATCAAGAACACA 60
DB 214 CACAAGTACACTACACTGCAATCATGCAAGATGCAACAAGCCAGATCAAGAACACA 273
QY 61 ACCCAACATACCTCACTAGATGATCTCTCAAGTGAATAGCTTCGCAATCTGTGAA 120
DB 274 ACCCAACATACCTCACTAGATGATCTCTCAAGTGAATAGCTTCGCAATCTGTGAA 333
QY 121 ATTACATGACAAACGACACCATATAGCTTCAACACACAGAGTCAAGTCAAACTG 180
DB 334 ATTACATGACAAACGACACCATATAGCTTCAACACACAGAGTCAAGTCAAACTG 393
QY 181 CAACCCACAAGTCAAGACTAAAAACACACACACCAACCAACCAAGCCAGCAAGCCC 240
DB 394 CAACCCACAAGTCAAGACTAAAAACACACACACCAACCAACCAAGCCAGCAAGCCC 453
QY 241 ACTACAAAACAAAGCCAAAACCAACCAACCAACCAATATGATTTTCACTTCGAA 300
DB 454 ACTACAAAACAAAGCCAAAACCAACCAACCAACCAATATGATTTTCACTTCGAA 513
QY 301 GTGTTAACTTTGTACCTGCGACATATGACAGCAACATTCACACTGCTGATCTGC 360
DB 514 GTGTTAACTTTGTACCTGCGACATATGACAGCAACATTCACACTGCTGATCTGC 573
QY 361 AAAAGATACCAACAAAAACAGGAAGAAACCAACCAAGCTTCAAAAAAACCA 420
DB 574 AAAAGATACCAACAAAAACAGGAAGAAACCAACCAAGCTTCAAAAAAACCA 633
QY 421 ACCTTCAAGACAAACAAAAAGATCTCAAACTCAAAACCAATTAACAGAGATACC 480
DB 634 ACCTTCAAGACAAACAAAAAGATCTCAAACTCAAAACCAATTAACAGAGATACC 693
QY 481 ACCACCAAGCCCAAGAGAGCAACATCAACACACCAACCAACCAATTAACAGAG 540
DB 694 ACCACCAAGCCCAAGAGAGCAACATCAACACACCAACCAACCAATTAACAGAG 753
QY 541 CTGCTCAGCAACACACAGAGAAATCCAAAATCACAAGTCAATGGAACCTTCCAC 600
DB 754 CTGCTCAGCAACACACAGAGAAATCCAAAATCACAAGTCAATGGAACCTTCCAC 813
QY 601 TCACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAGACCCA 660
DB 814 TCACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAGACCCA 873
QY 661 TCACACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAG 704
DB 874 TCACACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAG 917
RESULT 11

AR080424
LOCUS AR080424 894 bp DNA Linear PAT 31-AUG-2000
DEFINITION Sequence 28 from patent US 5968776.
ACCESSION AR080424
VERSION AR080424.1 GI:10007159
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 894)
AUTHORS Klein,M.H., Du,R.-P. and Ewasysbyn,M.E.
TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus
JOURNAL Patent: US 5968776-A 28 19-OCT-1999;
FEATURES
source location/Qualifiers
1..894
/organism="unknown"
BASE COUNT 364 a 288 c 93 g 149 t
ORIGIN
Query Match 97.3%; Score 696; DB 6; Length 894;
Best Local Similarity 100.0%; Pred. No. 2.7e-145;
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAGTACACTACACTGCAATCATGCAAGATGCAACAAGCCAGATCAAGAACACA 60
DB 199 CACAAGTACACTACACTGCAATCATGCAAGATGCAACAAGCCAGATCAAGAACACA 258
QY 61 ACCCAACATACCTCACTAGATGATCTCTCAAGTGAATAGCTTCGCAATCTGTGAA 120
DB 259 ACCCAACATACCTCACTAGATGATCTCTCAAGTGAATAGCTTCGCAATCTGTGAA 318
QY 121 ATTACATGACAAACGACACCATATAGCTTCAACACACAGAGTCAAGTCAAACTG 180
DB 319 ATTACATGACAAACGACACCATATAGCTTCAACACACAGAGTCAAGTCAAACTG 378
QY 181 CAACCCACAAGTCAAGACTAAAAACACACACACCAACCAACCAAGCCAGCAAGCCC 240
DB 379 CAACCCACAAGTCAAGACTAAAAACACACACACCAACCAACCAAGCCAGCAAGCCC 438
QY 241 ACTACAAAACAAAGCCAAAACCAACCAACCAACCAATATGATTTTCACTTCGAA 300
DB 439 ACTACAAAACAAAGCCAAAACCAACCAACCAACCAATATGATTTTCACTTCGAA 498
QY 301 GTGTTAACTTTGTACCTGCGACATATGACAGCAACATTCACACTGCTGATCTGC 360
DB 499 GTGTTAACTTTGTACCTGCGACATATGACAGCAACATTCACACTGCTGATCTGC 558
QY 361 AAAAGATACCAACAAAAACAGGAAGAAACCAACCAAGCTTCAAAAAAACCA 420
DB 559 AAAAGATACCAACAAAAACAGGAAGAAACCAACCAAGCTTCAAAAAAACCA 618
QY 421 ACCTTCAAGACAAACAAAAAGATCTCAAACTCAAAACCAATTAACAGAGATACC 480
DB 619 ACCTTCAAGACAAACAAAAAGATCTCAAACTCAAAACCAATTAACAGAGATACC 678
QY 481 ACCACCAAGCCCAAGAGAGCAACATCAACACACCAACCAACCAATTAACAGAG 540
DB 679 ACCACCAAGCCCAAGAGAGCAACATCAACACACCAACCAACCAATTAACAGAG 738
QY 541 CTGCTCAGCAACACACAGAGAAATCCAAAATCACAAGTCAATGGAACCTTCCAC 600
DB 739 CTGCTCAGCAACACACAGAGAAATCCAAAATCACAAGTCAATGGAACCTTCCAC 798
QY 601 TCACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAGACCCA 660
DB 799 TCACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAGACCCA 858
QY 661 TCACACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAG 696
DB 859 TCACACCTCTCTCGAAGGCAATCTAAAGCCCTTCTCAAGTCTCCACAACATCGAG 894

RESULT 12
 LOCUS AR092548 894 bp DNA linear PAT 08-SEP-2000
 DEFINITION Sequence 28 from patent us 5998169.
 ACCESSION AR092548
 VERSION AR092548.1 GI:10019302
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 894)
 AUTHORS Klein,M.H., Du,R.-P. and Ewaszyshyn,M.E.
 TITLE Multimeric hybrid gene encoding a chimeric protein which confers protection against parainfluenza virus and respiratory syncytial virus

JOURNAL Patent: US 5998169-A 28 07-DEC-1999;
 FEATURES Location/Qualifiers
 source 1..894
 BASE COUNT 364 a 288 c 93 g 149 t
 ORIGIN

Query Match 97.3%; Score 696; DB 6; Length 894;
 Best local similarity 100.0%; Pred. No. 2,7e-145;
 Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTGAACACTGCATCATCAAGATGCACAACGCCAGATCAAGAACACA 60
 DB 199 CACAAAGTCACACTGAACACTGCATCATCAAGATGCACAACGCCAGATCAAGAACACA 258
 QY 61 ACCCAACATACCTCAGTACAGATCCTCAGCTTGGAATGAGCTTCTCCAAATGTGTGAA 120
 DB 259 ACCCAACATACCTCAGTACAGATCCTCAGCTTGGAATGAGCTTCTCCAAATGTGTGAA 318
 QY 121 ATTACATCAACAACCAACCACTACTAGCTTCAACAACCAAGAGTCAAGTCAAACTG 180
 DB 319 ATTACATCAACAACCAACCACTACTAGCTTCAACAACCAAGAGTCAAGTCAAACTG 378
 QY 181 CAACCCCAACAGTCAAGACTAAACCAACAACCAACCAACCAACCAACCAACCAACGCC 240
 DB 379 CAACCCCAACAGTCAAGACTAAACCAACAACCAACCAACCAACCAACCAACCAACGCC 438
 QY 241 ACTACAAACCAACGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
 DB 439 ACTACAAACCAACGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498
 QY 301 GTGTTTAACCTTTAGCCCTGACGATATGACGACAAATTCCAACCTGCTGGCTATCTGC 360
 DB 499 GTGTTTAACCTTTAGCCCTGACGATATGACGACAAATTCCAACCTGCTGGCTATCTGC 558
 QY 361 AAAAGATATCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
 DB 559 AAAAGATATCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618
 QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
 DB 619 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678
 QY 481 ACCACCAAGCCCAAGAGGCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
 DB 679 ACCACCAAGCCCAAGAGGCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738
 QY 541 CTGCTCACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
 DB 739 CTGCTCACAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798
 QY 601 TCACACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACATCTCCAGCACCA 660
 DB 799 TCACACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACATCTCCAGCACCA 858
 QY 661 TCACACCTCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 696
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 859 TCACACCTCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 894

RESULT 13
 RSHICE 8510 bp ss-RNA linear VRL 29-NOV-2000
 LOCUS
 DEFINITION Human respiratory syncytial virus nonstructural protein (1C),
 nonstructural protein (1B), major nonnucleocapsid (N), phosphoprotein
 (P), protein (M), 1A (1A), G (G), protein (F) and
 envelope-associated protein (22K) gene, complete cds.
 M11486 K01459 K02719 K03348 K03349 M11217 M11244 M11487 M11505
 M11514 M11631 M12966
 M11486.1 GI:333925
 VERSION
 KEYWORDS envelope-associated protein; fusion glycoprotein; major
 nucleocapsid protein; major surface glycoprotein; matrix protein;
 nonstructural protein; phosphoprotein.
 SOURCE
 ORGANISM Human respiratory syncytial virus.
 REFERENCE 1 (bases 1 to 8510)
 AUTHORS Satake,M. and Venkatesan,S.
 TITLE Nucleotide sequence of the gene encoding respiratory syncytial
 virus matrix protein
 JOURNAL J. Virol. 50 (1), 92-99 (1984)
 MEDLINE 84138836
 PUBMED 6699948
 REFERENCE 2 (bases 2288 to 3191)
 AUTHORS Satake,M., Elango,N. and Venkatesan,S.
 TITLE Sequence analysis of the respiratory syncytial virus phosphoprotein
 gene
 JOURNAL J. Virol. 52 (3), 991-994 (1984)
 MEDLINE 85033973
 PUBMED 6548527
 REFERENCE 3 (bases 5602 to 7500)
 AUTHORS Collins,P.L., Huang,Y.T. and Wertz,G.W.
 TITLE Nucleotide sequence of the gene encoding the fusion (F)
 glycoprotein of human respiratory syncytial virus
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 81 (24), 7683-7687 (1984)
 MEDLINE 85088471
 PUBMED 6096849
 REFERENCE 4 (bases 7551 to 8510)
 AUTHORS Collins,P.L. and Wertz,G.W.
 TITLE The envelope-associated 22K protein of human respiratory syncytial
 virus: nucleotide sequence of the mRNA and a related polytranscript
 JOURNAL J. Virol. 54 (1), 65-71 (1985)
 MEDLINE 85135082
 PUBMED 3838351
 REFERENCE 5 (bases 8 to 1050; 7554 to 8506)
 AUTHORS Elango,N., Satake,M. and Venkatesan,S.
 TITLE mRNA sequence of three respiratory syncytial virus genes encoding
 two nonstructural proteins and a 22K structural protein
 JOURNAL J. Virol. 55 (1), 101-110 (1985)
 MEDLINE 85237684
 PUBMED 4009789
 REFERENCE 6 (bases 5602 to 7423)
 AUTHORS Elango,N., Satake,M., Colligan,J.E., Norrby,E., Camargo,E. and
 Venkatesan,S.
 TITLE Respiratory syncytial virus fusion glycoprotein: nucleotide
 sequence of mRNA, identification of cleavage activation site and
 amino acid sequence of N-terminus of F1 subunit
 JOURNAL Nucleic Acids Res. 13 (5), 1559-1574 (1985)
 MEDLINE 85215565
 PUBMED 2987829
 REFERENCE 7 (bases 4630 to 5543)
 AUTHORS Satake,M., Colligan,J.E., Elango,N., Norrby,E. and Venkatesan,S.

TITLE Respiratory syncytial virus envelope glycoprotein (G) has a novel structure
JOURNAL Nucleic Acids Res. 13 (21), 7795-7812 (1985)
MEDLINE 86067198
PUBMED 4069997
AUTHORS 9 (bases 4627 to 5544)
Wertz,G.W., Collins,P.L., Huang,Y., Gruber,C., Levine,S. and Ball,L.A.
TITLE Nucleotide sequence of the G protein gene of human respiratory syncytial virus reveals an unusual type of viral membrane protein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 82 (12), 4075-4079 (1985)
MEDLINE 85216636
PUBMED 3858865
AUTHORS 10 (bases 4173 to 4571)
Collins,P.L. and Wertz,G.W.
TITLE The 1A protein gene of human respiratory syncytial virus: nucleotide sequence of the mRNA and a related polycistronic transcript
JOURNAL Virology 141 (2), 283-291 (1985)
MEDLINE 86098645
PUBMED 3879976
AUTHORS 11 (bases 1 to 528; 552 to 1050)
Collins,P.L. and Wertz,G.W.
TITLE Nucleotide sequences of the 1B and 1C nonstructural protein mRNAs of human respiratory syncytial virus
JOURNAL Virology 143 (2), 442-451 (1985)
MEDLINE 86045905
PUBMED 2988021
AUTHORS 12 (bases 1081 to 2277)
Collins,P.L., Anderson,K., Langer,S.J. and Wertz,G.W.
TITLE Correct sequence for the major nucleocapsid protein mRNA of respiratory syncytial virus
JOURNAL Virology 146 (1), 69-77 (1985)
MEDLINE 85301974
PUBMED 3839952
AUTHORS 13 (bases 1051 to 1080; 2278 to 2287; 3192 to 3210; 4158 to 4172; 4572 to 4626; 5545 to 5601; 7501 to 7578)
Collins,P.L., Dickens,L.E., Buckler-White,A., Olmsted,R.A., Spriggs,M.K., Camarero,E. and Coelingh,K.V.
TITLE Nucleotide sequences for the gene junctions of human respiratory syncytial virus reveal distinctive features of intergenic structure and gene order
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 83 (13), 4594-4598 (1986)
MEDLINE 86259643
PUBMED 3460060
AUTHORS Reprint for [2] and clean copy sequence for [12],[5] kindly provided by P.Collins, 21-FEB-1986.
COMMENT Respiratory syncytial virus is a negative-strand RNA pleomorphic enveloped virus of the genus Pneumovirus. The matrix protein has no homology with the matrix proteins of other negative-stranded RNA viruses, implying that RS virus has undergone extensive evolutionary divergence. Two unidentified reading frames which potentially encode proteins were located: one overlaps the matrix protein and the other the F protein. The biological significance of these two reading frames is not clear. The positive strand is shown.

FEATURES
source location/Qualifiers
1..8510
/organism="Human respiratory syncytial virus"
/strain="A2"
/db_xref="taxon:11250"
/lab_host="HEp-2 cell"
1..528
/gene="1C"
1..528
/gene="1C"
55..474
/gene="1C"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAB59850.1"
/db_xref="GI:333926"
/translation="MGSNSLSMIVRLQNLFDNDEVALIKITCTYDKLIHLTNLALA

VIHITKLNGIVEVHTTSSDICPNNNIIVVKSNTFTMPVLQNGCYIWEMMELTHCSQPN
GLDNDNCEIKSKSKLSDSTMTNMYNQSLSELLGFDLNP"
171
/gene="1C"
/note="c in [11]; t in [6]"
/replace="c"
552..1050
/gene="1B"
552..1050
/gene="1B"
584..958
/gene="1B"
/codon_start=1
/product="nonstructural protein"
/protein_id="AAB59851.1"
/db_xref="GI:333927"
/translation="MDTTHNDNTQRLMITDMRPLSLFTITSITLRDITLHKTYLIN
HECIVRKDEQATFTFLVNTNEMKLLHKVGSTKYKRYTEYNTRYGTFPMPDIFTHDGF
LECGIKPTKHTPIIVKYDLPN"
747
/gene="1B"
/note="a in [11]; g in [6]"
/replace="g"
769
/gene="1B"
/note="g in [11]; a in [6]"
/replace="g"
1081..2277
/gene="N"
1081..2277
/gene="N"
1096..2271
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/protein_id="AAB59852.1"
/db_xref="GI:333928"
/translation="MALSKYKLVNDTLNKKDOLLSSSKYT.ORSYAGSDIDPNVYDOKHI
NKLCSGLITLEDANHKFTGLIGMLYAMSRIGREDTIKLTQBYHYVANGADVDTTRQ
DINKREKPFVLTLASLTLEIOTIEIESRSYKKMKLKEGEVAPRSTYPCGCMIT
LCIALVITKLAQDRSLGTAVIRANNVAKNEMKRYKGLIPDIANSFTEVEKERPH
FIDVFHFGLAQSTRGCSRVGEITFAGLEKNATGAGOVMLRWCVLAKSVANILGHAS
VQAEQEVVEVEYEAQRLGSDAGYVHLNPNKASLSLTQFPHFSSVYIGNAGLGLIM
GEGVGRNDLDVDAKAYAEOLKENGVINYSVLDTAELEAIRHQLNPKDMDEL"
2289..3196
/gene="P"
2289..3196
/gene="P"
2302..3027
/gene="P"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB59853.1"
/db_xref="GI:333929"
/translation="MEKFAPEFHGEDDANNRATFLESIKGKFTSPDPKKDSIISVN
SIDLEVTKESPITSNPIINPTNETDGTAGKNPNYORKEPVSEKEDPTSDKDFSKLY
KETLETFDNNSESSYSYEETINDQNTIARLDRIKESLITGLMHLTVVASAPT
SARDGIDAMIGLREEMIEKIRTEALMTNRLLEAMRLNRSEKMAKOTSDVSLNP
TSKRLNVLSEGNSDNDLSIEDP"
Query Match 91.6%; Score 654.8; DB 14; Length 8510;
Best Local Similarity 94.8%; Pred. No. 3; 6e-136;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

OY 1 CACAAAGTCACAGCAATGCAATCATCATGACGACCAACGCGATCAAGACACA 60
|||||
DB 4839 CACAAAGTCACAGCAATGCAATCATCATGACGACCAACGCGATCAAGACACA 4898
OY 61 ACCCAACATACCTCAGTCAAGATCTCAGCTTGATGAGCTTCCTCAATCTGTGAA 120
|||||
DB 4899 ACCCAACATACCTCAGTCAAGATCTCAGCTTGATGAGCTTCCTCAATCTGTGAA 4958
OY 121 ATTATATCACAAACCCACCATCTAGCTTCAACAACACCGAGATCAAGTCAAACTG 180

TITLE Direct Submission
JOURNAL Submitted (29-FEB-1996) RVS, LTD, NIAID, Bldg 7, Rm 118, 7 Center
Drive, MSC 0720, Bethesda, MD 20892, USA
FEATURES
source 1.15222
/organism="Human respiratory syncytial virus"
/strain="A2"
/db_xref="taxon:11250"
/note="mutant cp-RSV"
45..576
/gene="NS1 (1C)"
45..576
/gene="NS1 (1C)"
/product="nonstructural protein NS1"
99..518
/gene="NS1 (1C)"
/codon_start=1
/product="nonstructural protein NS1"
/protein_id="AAB86656.1"
/db_xref="GI:2627297"
/translation="MGSNSLSMKVRLQNLFDNDEVALIKITCYNDKLIHLTNALAKA
VIHTIKINGIVFVHVITSSDLCPPNNIVKSNFTMPYLQNGYIMEMELTCSQPN
GLDNCCEKFEKKSIDSMTVMYMOISLLECFDLP"
596..1098
/gene="NS2 (1B)"
596..1098
/gene="NS2 (1B)"
/product="nonstructural protein NS2"
628..1002
/gene="NS2 (1B)"
/codon_start=1
/product="nonstructural protein NS2"
/protein_id="AAB86657.1"
/db_xref="GI:2627298"
/translation="MDTTHNDNTPQRLMITMRPLSFTITSLTRDITLTKFIYLLN
HECIVKRLDERQATPTPLVNYEMKLHKVSGTKIKTYEYNTKGTGPPMPLFINHDF
LECIKPKTKHPITLYKYDLP"
1123..2327
/gene="N"
1125..2327
/gene="N"
/product="major nucleocapsid"
1140..2315
/gene="N"
/codon_start=1
/product="major nucleocapsid"
/protein_id="AAB86658.1"
/db_xref="GI:2627299"
/translation="MALSKVLTNTLTKDQLSSSKYITQSTGDSIDTPNYDYOKHI
NKLGMILITDANKRFTGLIGMLTAMSLGEREDTIKLRDAGYHVANGVDVTHRO
DINCKEMFEVLTLASLTTEIQINIEIESRSYKMKLKEGVAPEVRHSDPCGMII
LCIALVITKLAAQDRSGLTAVIRRANVULKNEMRYKGLPKDIANSFYVFGHPH
FIDVLFHFGIAQSDTRGSGRVEGIFAGLFNNAVGAGQMLRMGVLASIKIMIGHSS
VOAEQGVVEVEYEAOKLGEAGFYHILNNKASLISLTQPPHFSVVICAAALGIM
GEYGTGTPRNDLYAAKAYAEOLKENCYINISVLDITAELEAIKHQINPRDNDVEL"
2329..3242
/gene="P"
2329..3242
/gene="P"
/product="phosphoprotein"
2346..3071
/gene="P"
/codon_start=1
/product="phosphoprotein"
/protein_id="AAB86659.1"
/db_xref="GI:2627300"
/translation="MEKFAPEFHGEDANNRATKFLSINGKFTSPKDKKDSLIISVN
SIDLETFEKSTISNYSIINTNETDTPACKPKYQKRLPLVSFREDTTPSPDNPSKLI
KETLEEDNNNEESSYSYEBINDOTNNTTARDRIDEKUSEILIGMLHTLVASAGPT
SARDGITDANGLKEEMIEKIRTEALTNDRLEAMARLRNEESEKMAKDTSDVSLNP
TSEKLNLILEGNSDNDLSLEDF"
3252..4209
/gene="M"
3252..4209
/gene="M"

3252..4209
/gene="M"
/product="Matrix protein M"
3261..4031
/gene="M"
/codon_start=1
/product="Matrix protein M"
/protein_id="AAB86660.1"
/db_xref="GI:2627301"
/translation="METYVKNKLHGSTYTAAYQVNVLEKDDPASTITWPMQSSNP
ADLILKELANVILIKQISTPGPSLTVMANSSAVLACMPKFTTCANSLDERSKL
AYDVTPCEIKACSLTCLCKSKNMLTPTKIDLTMTNIDIIILCEPENVTSEKVIYI
PTYRSISVSRKNDLNTLENTTTEPKAINTAKIIPYSGLIYITVTDNKGAKRYIKP
OSQPIVDLGALEKESTIYYTTNKHATATPAIKPMED"
3970..4197
/gene="M"
/note="putative alternative translation of M gene"
/codon_start=1
/product="unknown"
/protein_id="AAB86661.1"
/db_xref="GI:2627302"
/translation="MLPQIGSTQLHDIQSPMKINLPLHQCYNYSKLSLYLHFTIT
ITNTLWPNQSKTYIKSQIILSHCLSDVLK"
4219..4628
/gene="SH (1A)"
4219..4628
/gene="SH (1A)"
/product="small hydrophobic protein"
4303..4497
/gene="SH (1A)"
/codon_start=1
/product="small hydrophobic protein"
/protein_id="AAB86662.1"
/db_xref="GI:2627303"
/translation="MENTSTIEPSSKFWPEFPLIMHTTIIISLIISIMAILNKL
CEYNVFNKTFELIPARAVNT"
4673..5595
/gene="G"
4673..5595
/gene="G"
/product="attachment G glycoprotein"
4688..5584
/gene="G"
/codon_start=1
/product="attachment G glycoprotein"
/protein_id="AAB86663.1"
/db_xref="GI:2627304"
/translation="MSKNKQDQRAKTLERTWDTLNLHLFTSSCYKLNLSVAQITLS
ILMILISTSLIAAIFFIASANRKYPTPAIIODATSOIKNTPTVLTQNPOLGISPS
NPSEILISQITTLASTTPGVKSTLQSTYVTKNTTQTPQSPKTKOKONKPPSKPN
NDHFEVFNVPSCISNNPTCAICRKRIPNKKPKKTKTKPKPTLTKTKKDPKO
TTSKKEVTPKPREPTINTKNTITTLTISNTGNPELTQMETFHTSSGQNSP
SOVSTTSEVSSQSSPPNTPRO"
5648..7550
/gene="F"
5648..7550
/gene="F"
/product="fusion F glycoprotein precursor"
5661..7385
/gene="F"
/codon_start=1
/product="fusion F glycoprotein precursor"
/protein_id="AAB86664.1"
/db_xref="GI:2627305"
/translation="MELILKANALITTLTAVTFCASGQNTIEEYQSTCSAVSKGY
LSLRKQWTSVTTTLLSNIKENKKGKGTAKKLLIKOEIDLRKNVATELIQILMSTPA
TNKARRELPFRMNTLNNAKTNTVLSKRRKRLGLGVGSATASAVASKVLHL
EGEVRKIKSALLSTNKAAVLSLVNSVLSKVLIDKNTYIDKQLPLPVNQSQSSINIA
TVTEFOQKNRRLIETIRERSVNAVGYTTPVSTYMLNSELISLINDPITNDQKILMSN
NVOIVQOQSYISINSIKEBVLAVVOLPLYGVIDPCKMLHTSPICTTNTRKGSNCL
TRDRGMVCDNAGSVSFPPQAECKVQSNRVCDDMNSLTLPEVNLINVDIFNKYD
CKLMPSTKTPVSSVLTISGAIYSCYGTCTGASNNKRGILIKTFSSGCDVSNKGYDTV
SGNTLYYVKNQDQKSLYVGEPIINFTDPLVFPSEDFASISOVNEKINQSLATIRK

SDLLHNYNACKSTINIMITIIIVIIILLSLAVGLLLYCKARSPVTLRSKQSLG
INNTAFSN"
7597..8557
/gene="M2 (22K)"
7597..8557
/gene="M2 (22K)"
/product="matrix protein M2"
7606..8190
/gene="M2 (22K)"
/codon_start=1
/product="matrix protein M2"
/protein_id="AAB8665.1"
/db_xref="GI:2627306"
/translation="MSRRNPCKFEIRGHCLNGKRHPSHNYFEMPYPHLLYRONFMLN
RLIKSMKSIDTLEISGAALDREELGVGLSYIGSINNTIKOSACVMSKL
LTELNSDDIKLRONEELNSPKIRYNTVISYIESNRKNKQTHLLKRLPADYLAKT
IKNLDLHKSTFINNPKESTVSDTNDAKNDDT"
8159..8431
/gene="M2 (22K)"
/note="putative alternative translation of M2 gene"
/codon_start=1
/product="unknown"
/protein_id="AAB8666.1"
/db_xref="GI:2627307"
/translation="WTMPKIMILPDKYPGSIYSILITSRCRVIMYNOKNTLYFNQNNP

Query Match 91.6%; Score 654.8; DB 14; Length 15222;
Best Local Similarity 94.8%; Pred. No. 3.4e-136;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACATACTGCAATCATCAAGATGCAACAAGCCAGATCAAGAACACA 60
DB 4886 CACAAGTCACACCACTGCAATCATCAAGATGCAACAAGCCAGATCAAGAACACA 4945
QY 61 ACCCCACATFACCTCAGATCTCTGAGCTTGGAAATCAGCTTCTCAATCTGTCTGAA 120
DB 4946 ACCCCACATFACCTCAGATCTCTGAGCTTGGAAATCAGCTTCTCAATCTGTCTGAA 5005
QY 121 ATTACATCAAAACCAACCACTAGCTTCAACAACCAAGAGTCAAGTCAAACTG 180
DB 5006 ATTACATCAAAACCAACCACTAGCTTCAACAACCAAGAGTCAAGTCAAACTG 5065
QY 181 CAACCCACAACAGTCAAGACTTAAACACACAAACCAACCAACCAACCAAGCCG 240
DB 5066 CAATCCACAACAGTCAAGACTTAAACACACAAACCAACCAACCAACCAAGCCG 5125
QY 241 ACTACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 5126 ACCACAAAACAGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5185
QY 301 GTGTTAACTTGTACCTGACATATGACAAATCAATCAACCTGCTGGCTATCTGC 360
DB 5186 GTGTTAACTTGTACCTGACATATGACAAATCAATCAACCTGCTGGCTATCTGC 5245
QY 361 AAAAGAAATCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 420
DB 5246 AAAAGAAATCAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5305
QY 421 ACCTTAAGACACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAAAGTACCC 480
DB 5306 ACCTTAAGACACCAAAAGATCTCAAACTCAAAACCACTAAACCAAGAAAGTACCC 5365
QY 481 ACCACCAAGCCACAGAAAGCCAAACATCAACACCAACCAACCAACCAACCAACCA 540
DB 5366 ACCACCAAGCCACAGAAAGCCAAACATCAACACCAACCAACCAACCAACCAACCA 5425
QY 541 CTGCTACCAACACACACAGAAATCCAAATCTCAAGTCAATGGAACCTTCAC 600
DB 5426 CTGCTACCAACACACACAGAAATCCAAATCTCAAGTCAATGGAACCTTCAC 5485
QY 601 TCAACCTCTCGAAGGCAATTAAGCCCTTCAAGTCTCCACACATCCGACACCA 660
DB 5486 TCAACCTCTCGAAGGCAATTAAGCCCTTCAAGTCTCTACACATCCGAGTACCCA 5545

QY 661 TCACACCTTCATCTCCACCAACACACAGCCAGTAGTATTAAAAA 714
DB 5546 TCACACCTTCATCTCCACCAACACACAGCCAGTAGTATTAAAAA 5599

Search completed: May 1, 2003, 11:59:54
Job time : 1461.68 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 07:53:37 ; Search time 139.502 seconds

(without alignments)
11542.374 Million cell updates/sec

Title: US-09-462-816-3

Perfect score: 715
Sequence: 1 cacaaagtcacactaacac.....gtagtattacaaaaaaaaa 715

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*

```
1: /SIDS2/gcgdata/geneq/geneq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneq/geneq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneq/geneq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneq/geneq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneq/geneq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneq/geneq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneq/geneq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneq/geneq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneq/geneq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneq/geneq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneq/geneq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneq/geneq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneq/geneq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneq/geneq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneq/geneq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneq/geneq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneq/geneq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneq/geneq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneq/geneq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneq/geneq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneq/geneq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneq/geneq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneq/geneq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneq/geneq-emb1/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match	Length	DB ID	Description
1	715	100.0	715	20	AA08422
2	715	100.0	920	14	AA045686
3	715	100.0	920	20	AA08421
4	659	92.2	935	8	AA070784
5	659	92.2	935	19	AA018736
6	657.4	91.6	935	13	AA029623
7	654.8	91.6	15210	20	AA059703
8	654.8	91.6	15222	18	AA078440
9	654.8	91.6	15223	18	AA063430

10	654.8	91.6	15223	19	AA017553	Respiratory syncyt
11	654.8	91.6	15223	21	AA086743	Respiratory syncyt
12	649.6	90.9	918	22	AA088494	Human RSV G-protei
13	646.2	90.4	897	24	AA020145	Respiratory syncyt
14	445.8	62.3	696	19	AA038298	Respiratory syncyt
15	300.4	42.0	15225	19	AA017552	Respiratory syncyt
16	300.4	42.0	15225	21	AA088744	Respiratory syncyt
17	300.4	42.0	15225	24	AA025520	Human respiratory
18	295.6	41.3	15229	19	AA018276	Human respiratory
19	295.6	41.3	15229	20	AA022910	Nucleotide sequenc
20	295.6	41.3	15229	20	AA035268	DNA encoding the L
21	273.6	38.3	15218	19	AA018275	RSV isolate 28 w11
22	273.6	38.3	15218	20	AA022909	Nucleotide sequenc
23	273.6	38.3	15218	20	AA022914	Nucleotide sequenc
24	273.6	38.3	15218	20	AA035267	DNA encoding the L
25	273.6	38.3	15219	19	AA018277	RSV vaccine 2833F
26	273.6	38.3	15219	19	AA018278	RSV vaccine 2820L
27	273.6	38.3	15219	19	AA018279	RSV revertant 2833
28	273.6	38.3	15219	19	AA018280	RSV revertant 2820
29	273.6	38.3	15219	20	AA022911	Nucleotide sequenc
30	273.6	38.3	15219	20	AA022912	Nucleotide sequenc
31	273.6	38.3	15219	20	AA022913	Nucleotide sequenc
32	273.6	38.3	15219	20	AA035269	DNA encoding the L
33	273.6	38.3	15219	20	AA035270	DNA encoding the L
34	273.6	38.3	15219	20	AA035271	DNA encoding the L
35	273.6	38.3	15219	20	AA035272	DNA encoding the L
36	212.8	29.8	1050	17	AA013647	Incodes Streptococ
37	212.8	29.8	1050	22	AA047111	Nucleotide sequenc
38	212.8	29.8	1356	22	AA080153	RSV subgroup A clo
39	210.8	29.5	303	16	AA030486	RSV subgroup A prote
40	210.8	29.5	303	17	AA027066	Respiratory Syncyt
41	210.8	29.5	303	17	AA027073	RSV G protein anti
42	210.8	29.5	303	17	AA013649	RSV G protein anti
43	210.8	29.5	303	20	AA030478	Nucleotide sequenc
44	210.8	29.5	303	21	AA058882	
45	210.8	29.5	303	22	AA078459	

ALIGNMENTS

RESULT 1	
ID	AA08422 standard; cDNA; 715 BP.
XX	XX
AC	AA08422:
XX	XX
XX	28-JUN-1999 (first entry)
XX	XX
DE	G protein gene fragment of respiratory syncytial virus.
XX	XX
KW	G protein: respiratory syncytial virus; RSV; recombinant vector;
KW	vacine; immune response; immunogenicity; tPA; antibody;
KW	tissue plasminogen activator; ss.
OS	Respiratory syncytial virus (RSV).
XX	XX
FT	Key
FT	Location/Qualifiers
FT	1..702
FT	/tag= a
FT	/product= "Secreted G protein"
PN	W09904010-A1.
XX	XX
PD	28-JAN-1999.
XX	XX
PF	16-JUL-1998; 98MO-CA00697.
XX	XX
PR	18-JUL-1997; 97US-0896442.
XX	XX
PA	(CONN-) CONNAUGHT LAB LTD.
XX	XX
PI	Klein MH, Li X, Sambhara S;

```

XX  WPI: 1999-132254/11.
DR  P-PSDB: AAW96314.
XX
PT  Immunogenic composition for generating antibodies against
PT  respiratory syncytial virus - comprises non-replicating vector
PT  containing the protein G sequence, useful in protective vaccines and
PT  to raise antibodies for diagnosis
XX
PS  Claim 8; Figure 3; 67pp; English.
XX
CC  The respiratory syncytial virus (RSV) G protein can be used in
CC  vaccines by inserting the G protein gene into a non-replicating
CC  vector. The G protein is placed under the control of alternative
CC  signal and expression sequences, for example the chimeric G protein
CC  produced may also comprise the signal peptide of tissue plasminogen
CC  activator (tPA). The recombinant vector may also comprise sequences
CC  upstream of the G protein gene which enhance the G proteins
CC  immunoprotective ability. The resulting immunogenic composition will
CC  generate antibodies directed against the RSV G protein when
CC  administered to a host organism. The composition is useful as a
CC  vaccine to immunise against RSV-associated disease, particularly
CC  resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC  by usual immunisation and cell fusion methods.
XX
SQ  Sequence 715 BP; 305 A; 240 C; 70 G; 100 T; 0 other;
Query Match      100.0%; Score 715; DB 20; Length 715;
Best Local Similarity 100.0%; Pred. No. 3.2e-162;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0Y  1 CACAAAGTCACACTAACAAGTCATATCAAGATGCAACAGCCAGATCAAGACACA 60
DB  1 CACAAAGTCACACTAACAAGTCATATCAAGATGCAACAGCCAGATCAAGACACA 60
0Y  61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 120
DB  61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 120
0Y  121 ATTACTGCAACCAACGACCTACTGATGCTTCAACACAGCGAGGTCAAGTCAAGT 180
DB  121 ATTACTGCAACCAACGACCTACTGATGCTTCAACACAGCGAGGTCAAGTCAAGT 180
0Y  121 ATTACTGCAACCAACGACCTACTGATGCTTCAACACAGCGAGGTCAAGTCAAGT 180
DB  121 ATTACTGCAACCAACGACCTACTGATGCTTCAACACAGCGAGGTCAAGTCAAGT 180
0Y  181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240
DB  181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 240
0Y  241 ACTACAAACCAACGACCTACTGATGCTTCAACACAGCGAGGTCAAGTCAAGT 300
DB  241 ACTACAAACCAACGACCTACTGATGCTTCAACACAGCGAGGTCAAGTCAAGT 300
0Y  301 GTGTTTAAGCTTTGACCTGACGATATGACGACCAATGCAAGTCTGAGTCTGAG 360
DB  301 GTGTTTAAGCTTTGACCTGACGATATGACGACCAATGCAAGTCTGAGTCTGAG 360
0Y  361 AAAAGAATACCAAAACCAAGGAGAAAGAAACCAACCAAGTCTGAGTCTGAGT 420
DB  361 AAAAGAATACCAAAACCAAGGAGAAAGAAACCAACCAAGTCTGAGTCTGAGT 420
0Y  421 ACCTTAAAGACCAAAACCAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 480
DB  421 ACCTTAAAGACCAAAACCAAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 480
0Y  481 ACCACCAAGCCACAGAGGACCAACATCAACACCAACCAACCAACATCAACATCA 540
DB  481 ACCACCAAGCCACAGAGGACCAACATCAACACCAACCAACCAACATCAACATCA 540
0Y  541 CTGCTACCAACACACACAGGAAATCAAAATCAAGTCAAGTCAAGTCAAGTCAAG 600
DB  541 CTGCTACCAACACACACAGGAAATCAAAATCAAGTCAAGTCAAGTCAAGTCAAG 600
0Y  601 TCAACCTCGCGAAGGACATCAAGCCCTTCAAGTCTGAGTCTGAGTCTGAGTCA 660
DB  601 TCAACCTCGCGAAGGACATCAAGCCCTTCAAGTCTGAGTCTGAGTCTGAGTCA 660

```

```

DB  601 TCAACCTCGCGAAGGACATCAAGCCCTTCAAGTCTGAGTCTGAGTCTGAGTCA 660
0Y  661 TCACAAACCTCATCTCCACCAACACACAGCGAGTCTGATTAATAAAAAAAAA 715
DB  661 TCACAAACCTCATCTCCACCAACACACAGCGAGTCTGATTAATAAAAAAAAA 715

RESULT 2
AAQ45686
ID  AAQ45686 standard; DNA; 920 BP.
XX
AC  AAQ45686;
XX
DT  13-JAN-1994 (first entry)
XX
DE  Respiratory syncytial virus (RSV) G gene.
XX
KW  PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine;
XX  ds.
OS  Respiratory syncytial virus.
XX
FH  Key Location/Qualifiers
FT  CDS 8..901
FT  /*tag= a
FT  /product= RSV G protein
FT  misc_feature 89..176
FT  /*tag= b
FT  /label= Transmembrane anchor domain
XX
PN  W09314207-A.
XX
PD  22-JUL-1993.
XX
PE  05-JAN-1993; 93MO-CA00001.
XX
PR  06-JAN-1992; 92GB-0000117.
XX
PA  (CONN-) CONNADGHT LAB LTD.
XX
PI  Ewasayhyn ME, Klein MH;
XX
DR  WPI: 1993-243222/30.
DR  P-PSDB: AAR39286.
XX
PT  Multimeric hybrid genes and their chimeric proteins - are
PT  vaccines against multiple pathogenic infections e.g.
PT  para-influenza virus and respiratory syncytial virus
XX
PS  Claim 11; Figure 7A-7D; 80pp; English.
XX
CC  A novel multimeric hybrid gene is used as a vaccine. The gene
CC  consists of two gene sequences which are linked and encode antigenic
CC  regions, these two sequences being derived from two different
CC  pathogens (para-influenza virus (PIV) and respiratory syncytial virus
CC  (RSV)). The gene sequences that are particularly used are those
CC  which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and
CC  G proteins (AAQ45685, AAQ45686).
XX
SQ  Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;
Query Match      100.0%; Score 715; DB 14; Length 920;
Best Local Similarity 100.0%; Pred. No. 3.4e-162;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
0Y  1 CACAAAGTCACACTAACAAGTCATATCAAGATGCAACAGCCAGATCAAGACACA 60
DB  206 CACAAAGTCACACTAACAAGTCATATCAAGATGCAACAGCCAGATCAAGACACA 265
0Y  61 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 120
DB  266 ACCCAACATACCTCAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGT 325

```

QY 121 ATTACATCAAAACCAACCATCTAGCTTCACACAGCAGAGTCAAGTCAACCTG 180
DB 326 ATTACATCAAAACCAACCATCTAGCTTCACACAGCAGAGTCAAGTCAACCTG 385
QY 181 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAACCAAGCCC 240
DB 386 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAACCAAGCCC 445
QY 241 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTAACTTTGACCTGACGATATGACGCAATTCACCTGCTGGCTATCTGC 360
DB 506 GTGTTAACTTTGACCTGACGATATGACGCAATTCACCTGCTGGCTATCTGC 565
QY 361 AAAAATAATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAATAATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685
QY 481 ACCACCAAGCCCAAGAGGCAACATCAACACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCCAAGAGGCAACATCAACACCAACCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTCAGCAACACACCAAGGAATGCCAAATCTCAATGCAATGGAACCTTCAC 600
DB 746 CTGCTCAGCAACACACCAAGGAATGCCAAATCTCAATGCAATGGAACCTTCAC 805
QY 601 TCACACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTGCACACATCCGAGACCCA 660
DB 806 TCACACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTGCACACATCCGAGACCCA 865
QY 661 TCACACCTCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 866 TCACACCTCTCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920

RESULT 3
AA08421

ID AA08421 standard; cDNA: 920 BP.

AC AA08421:

DT 28-JUN-1999 (first entry)

DE Membrane bound G protein gene of respiratory syncytial virus.

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW tissue plasminogen activator; ss.

OS Respiratory syncytial virus (RSV).

FH Key location/Qualifiers

FT CDS 8..904 /tag= a /product= "Membrane bound G protein"

PN W09904010-A1.

PD 28-JAN-1999.

PF 16-JUL-1998; 98WO-CA00697.

PR 18-JUL-1997; 97US-0896442.

PA (CONN-) CONNBUGHT LAB LTD.

XX Klein MH, Li X, Sambhara S;
PI

XX WI: 1999-132254/11.
DR P-PSDB; AAM96313.
XX Immunogenic composition for generating antibodies against
PT respiratory syncytial virus - comprises non-replicating vector
PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis
PS Claim 3; Figure 2; 67pp; English.
XX The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternative
CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods.
XX Sequence 920 BP; 380 A; 290 C; 95 G; 155 T; 0 other;

Query Match 100.0%; Score 715; DB 20; Length 920;

Best Local Similarity 100.0%; Pred. No. 3,4e-162;

Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCAAC 60
DB 206 CACAAAGTCAACATCAACATCAACATCAACATCAACATCAACATCAACATCAACATCAAC 265
QY 61 ACCCAACATACCTCTAGAGATCTGAGCTTGGATGAGTCTGCAATCTGCTGAA 120
DB 266 ACCCAACATACCTCTAGAGATCTGAGCTTGGATGAGTCTGCAATCTGCTGAA 325
QY 121 ATTACATCAAAACCAACCAACATCAACATCAACATCAACATCAACATCAACATCAAC 180
DB 326 ATTACATCAAAACCAACCAACATCAACATCAACATCAACATCAACATCAACATCAAC 385
QY 181 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAACCAACCAACCAACCA 240
DB 386 CAACCCACACAGTCAAGCTTAAAAACACACACACCAACCAACCAACCAACCAACCAACCA 445
QY 241 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAAACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTAACTTTGACCTGACGATATGACGCAATTCACCTGCTGGCTATCTGC 360
DB 506 GTGTTAACTTTGACCTGACGATATGACGCAATTCACCTGCTGGCTATCTGC 565
QY 361 AAAAATAATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAATAATTCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685
QY 481 ACCACCAAGCCCAAGAGGCAACATCAACACCAACCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCCAAGAGGCAACATCAACACCAACCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTCAGCAACACACCAAGGAATGCCAAATCTCAATGCAATGGAACCTTCAC 600
DB 746 CTGCTCAGCAACACACCAAGGAATGCCAAATCTCAATGCAATGGAACCTTCAC 805
QY 601 TCACACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTGCACACATCCGAGACCCA 660
|||||

```

Db      806  TCAACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACAACATCCGACACCA 865
QY      661  TCACAACTCATCTCCACCCACACACAGCCAGTATTATTAACAAAAA 715
      |||
Db      866  TCACAACTCATCTCCACCCACACACAGCCAGTATTATTAACAAAAA 920
      |||

RESULT 4
AAV18784
ID      AAV18784 standard; cDNA; 935 BP.
XX
AC      AAV18784:
XX
DT      05-APR-1991 (first entry)
XX
DE      Sequence encoding human respiratory syncytial virus (HRSV) A2 strain
DE      G protein.
XX
KW      Vaccine; ss.
XX
OS      Human respiratory syncytial virus (HRSV).
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      16..913
FT      /tag= a

XX      MO8704185-A.
XX      16-JUL-1987.
XX
PF      23-DEC-1986; 86MO-US02756.
XX
PR      14-JAN-1986; 86US-0818740.
XX
PA      (UYN(-) UNIV OF N CAROLINA.
PA      (WERT/) WERTZ G W.
PA      (WERT/) WERTZ G W.
XX
DR      WPI; 1987-206300/29.
DR      P-PSDB; AAP70845.
XX
PT      Vaccines for human respiratory virus - comprising proteins or
PT      fragment encoded by a DNA sequence coding for human respiratory
PT      syncytial virus proteins.
XX
PS      Disclosure; Chart 13; 57pp; English.
XX
CC      A novel plasmid which comprises a DNA sequence encoding this
CC      protein, and the protein itself, are claimed, for use as HRSV
CC      vaccines. The vaccine can be administered to pregnant women or to
CC      women of child bearing age to stimulate maternal antibodies.
CC      CC
CC      Infants can also be vaccinated at 2-3 months of age.
XX
SO      Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match      92.2%; Score 659; DB 8; Length 935;
Best Local Similarity 95.1%; Pred. No. 9.5e-149;
Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY      1  CACAAAGTCACTAAAGTCAATCATAGATGCAAGCAAGCCAGATCAAGAACACA 60
      |||
Db      214  CACAAAGTCACTAAAGTCAATCATAGATGCAAGCAAGCCAGATCAAGAACACA 273
      |||
QY      61  ACCCAACATACCTCACTAGATCCGACCTTGAATCAGCTTCTCAATCTGTGAA 120
      |||
Db      274  ACCCAACATACCTCACTAGATCCGACCTTGAATCAGCTTCTCAATCTGTGAA 333
      |||
QY      121  ATTACATGACAAACACCACTACTAGTTCACAAACACCAAGAGTCAAGTAAACCTG 180
      |||
Db      334  ATTACATGACAAACACCACTACTAGTTCACAAACACCAAGAGTCAAGTAAACCTG 393
      |||
QY      181  CAACCCACACAGTCAAGACTAAACACAAACCAACCAACACCAAGCAGCC 240
      |||

```

```

Db      394  CAATCACAACAGTCAAGACCAAAAAACAAACAACTCAAAACACCAAGCAGCC 453
QY      241  ACTACAAACAAAGCCAAACAAACCAACCAACCAATTAATGATTTCACTTGAA 300
      |||
Db      454  ACCACAAACAAAGCCAAACAAACCAACCAACCAATTAATGATTTCACTTGAA 513
      |||
QY      301  GTGTTAACTTTGTACCTGACGATATGACAAACAACTGCTGGGCTATCTGC 360
      |||
Db      514  GTGTTCAACTTTGTACCTGACGATATGACAAACAACTGCTGGGCTATCTGC 573
      |||
QY      361  AAAAGATATCCAAACAAACCAAGAAAGAAACCAAGCCGCTACAAAAACCA 420
      |||
Db      574  AAAAGATATCCAAACAAACCAAGAAAGAAACCAAGCCGCTACAAAAACCA 633
      |||
QY      421  ACCCTCAAGACCAACCAAAAGATCTCAACCTCAAAACCACTAAACCAAGAGTACC 480
      |||
Db      634  ACCCTCAAGACCAACCAAAAGATCTCAACCTCAAAACCACTAAATCAAGAGTACC 693
      |||
QY      481  ACCACCAAGCCCAAGAGCCAAACATCAACACCAACCAACCAACATCAACTACA 540
      |||
Db      694  ACCACCAAGCCCAAGAGCCAAACATCAACACCAACCAACCAACATCAACTACA 753
      |||
QY      541  CTGCTCACCACCAACACCAAGGAAATCCAAACCTCAACAGTCAAAATGGAACCTTCAC 600
      |||
Db      754  CTACTCACCCTCCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCA 813
      |||
QY      601  TCACACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCA 660
      |||
Db      814  TCACACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACATCCCA 873
      |||
QY      661  TCACAACTCATCTCCACCCACACACAGCCAGTATTATTAACAAAAA 715
      |||
Db      874  TCACAACTCATCTCCACCCACACACAGCCAGTATTATTAACAAAAA 928
      |||

RESULT 5
AAV18736
ID      AAV18736 standard; cDNA; 935 BP.
XX
AC      AAV18736;
XX
DT      11-JUN-1998 (first entry)
XX
DE      HRSV glycoprotein G cDNA.
XX
KW      HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine; ss.
XX
OS      Human respiratory syncytial virus.
XX
FH      Key
FT      CDS
FT      Location/Qualifiers
FT      16..912
FT      /tag= a
FT      /product= glycoprotein_G

XX      US5716823-A.
XX
PD      10-FEB-1998.
XX
PF      12-MAY-1997; 97US-0854783.
XX
PR      13-JUL-1988; 88US-0218737.
PR      14-JAN-1986; 86US-0818740.
PR      23-DEC-1986; 86MO-US02756.
PR      11-JUN-1992; 92US-0897171.
PR      12-MAY-1997; 97US-0854783.
XX
PA      (PHRA ) PHARMACIA & UPJOHN CO.
XX
PI      Collins PL; Wertz GW;
XX
DR      WPI; 1998-144802/13.
DR      P-PSDB; AAV47605.
XX

```


PT Production of human respiratory syncytial virus glyco-protein F or G
 PT - by culturing eukaryotic host cells transfected with corresponding
 PT DNA

XX Example 1; Columns 27-28; 17pp; English.

XX The present sequence was used in the development of a novel method
 CC for the production of human respiratory syncytial virus (HRSV)
 CC glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
 CC culturing eukaryotic host cells transfected with an isolated DNA
 CC sequence encoding HRSV gpF or gpG. The gp can be used to prepare
 CC vaccines against HRSV.

XX Sequence 935 BP; 383 A; 293 C; 100 G; 159 T; 0 other;

Query Match 92.2%; Score 659; DB 19; Length 935;
 Best Local Similarity 95.1%; Pred. No. 9,5e-149;
 Matches 680; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 1 CACAAGTCACTAACAACCTCAATATCAAGATGCAACAGCCAGATCAAGAACACA 60
 DB 214 CACAAGTCACTAACAACCTCAATATCAAGATGCAACAGCCAGATCAAGAACACA 273
 QY 61 ACCCCAATACCTCACTAGATGATGCTTGAATGATGCTTCAATCTGTCTGAA 120
 DB 274 ACCCCAATACCTCACTAGATGATGCTTGAATGATGCTTCAATCTGTCTGAA 333
 QY 121 ATTACATCAACAACACCCATAGCTTCAACACAGGAGTCAAGTCAAACTCG 180
 DB 334 ATTACATCAACAACACCCATAGCTTCAACACAGGAGTCAAGTCAAACTCG 393
 QY 181 CAACCCACACAGTCAAGCTTAAACACACACACACCCAAACCAACCCAGAACGCC 240
 DB 394 CAATCCACACAGTCAAGCTTAAACACACACACACCCAAACCAACCCAGAACGCC 453
 QY 241 ACTACAAACACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
 DB 454 ACCAACAACACAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 513
 QY 301 GTGTTAACTTGTACCTGACGATATGACGACCAATTCACCTGCTGCTATCTGC 360
 DB 514 GTGTTAACTTGTACCTGACGATATGACGACCAATTCACCTGCTGCTATCTGC 573
 QY 361 AAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
 DB 574 AAAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
 QY 421 ACCTTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
 DB 634 ACCTTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 693
 QY 481 ACCACCAAGCCCAAGAGACCAACCATCAACACACCAACCAACCAACCAACCA 540
 DB 694 ACCACCAAGCCCAAGAGACCAACCATCAACACACCAACCAACCAACCAACCA 753
 QY 541 CTGCTCAGCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
 DB 754 CTGCTCAGCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 813
 QY 601 TCAACCTCTCCGAAAGCAATCTTAAGCCTTCTCAAGTCTCAACACATCCGAGACGCCA 660
 DB 814 TCAACCTCTCCGAAAGCAATCTTAAGCCTTCTCAAGTCTCAACACATCCGAGACGCCA 873
 QY 661 TCACACCCCTGATCTCCACCCACACACACAGCAGTAGTATTAAAAAAGAAAA 715
 DB 874 TCACACCCCTGATCTCCACCCACACACACAGCAGTAGTATTAAAAAAGAAAA 928

RESULT 6
 AAQ29623
 ID AAQ29623 standard; DNA; 935 BP.
 XX
 AC AAQ29623;

XX 03-MAR-1993 (first entry)
 DT
 XX HSRV glycoprotein G (gpG).
 DE
 XX
 KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;
 KW major capsid protein; N; ss.
 OS Human respiratory syncytial virus strain A2.
 FH
 FH Key Location/Qualifiers
 FT CDS 16..912
 FT /*tag= a
 FT /label= G_protein
 FT misc_feature 16..22
 FT /*tag= b
 FT /note= "oligonucleotide used to probe for
 FT full length cDNA"
 FT misc_feature 898..912
 FT /*tag= c
 FT /note= "oligonucleotide used to specifically
 FT prime the reverse transcription reaction
 FT for making the first strand of the cDNA"
 PD US5149650-A.
 PD 22-SEP-1992.
 PE 14-JAN-1986; 86US-0818740.
 PR 14-JAN-1986; 86US-0818740.
 PR 13-JUL-1988; 88US-0218737.
 PA (UNCL)- UNIV NORTH CAROLINA.
 XX
 XX Collins PL, Wertz GW;
 XX
 XX WPI: 1992-340247/41.
 DR P-PSDB; AAR25302.
 DR
 PT Vaccines for human respiratory virus - include structural genes
 PT coding for native structural viral proteins and immunogenic
 PT fragments
 PT
 PS Disclosure; Page 18; 21pp; English.
 XX
 CC The sequences of mRNA encoding HRSV structural proteins are given in
 CC AAQ29622-26. The proteins are F, G, 22K, 9.5K and major capsid
 CC protein N. The sequences and encoded proteins are useful for
 CC preparing vaccines against HRSV. The vaccines can be used to confer
 CC immunity against respiratory tract infections on human subjects.
 CC
 QY Sequence 935 BP; 382 A; 294 C; 100 G; 159 T; 0 other;
 SO
 Query Match 91.9%; Score 657.4; DB 13; Length 935;
 Best Local Similarity 95.0%; Pred. No. 2.3e-148;
 Matches 679; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 CACAAGTCACTAACAACCTCAATATCAAGATGCAACAGCCAGATCAAGAACACA 60
 DB 214 CACAAGTCACTAACAACCTCAATATCAAGATGCAACAGCCAGATCAAGAACACA 273
 QY 61 ACCCCAATACCTCACTAGATGATGCTTGAATGATGCTTCAATCTGTCTGAA 120
 DB 274 ACCCCAATACCTCACTAGATGATGCTTGAATGATGCTTCAATCTGTCTGAA 333
 QY 121 ATTACATCAACAACACCCATAGCTTCAACACAGGAGTCAAGTCAAACTCG 180
 DB 334 ATTACATCAACAACACCCATAGCTTCAACACAGGAGTCAAGTCAAACTCG 393
 QY 181 CAACCCACACAGTCAAGCTTAAACACACACACACCCAAACCAACCCAGAACGCC 240
 DB 394 CAATCCACACAGTCAAGCTTAAACACACACACACCCAAACCAACCCAGAACGCC 453


```

OY 1 CACAAGTCACTACATGCAATTCATACAGATGCAACAAGCCAGATCAAGAACA 60
   |||||||
Db 4887 CACAAGTCACTACATGCAATTCATACAGATGCAACAAGCCAGATCAAGAACA 4946
OY 61 ACCCAACATACCTCACTAGATGCTCTGAGTTCAGTCTTCCATCTGTCTGAA 120
   |||||||
Db 4947 ACCCAACATACCTCACTAGATGCTCTGAGTTCAGTCTTCCATCTGTCTGAA 5006
OY 121 ATTACATCAACCAACCACTACTAGTTCACACACAGGAGTCAAGTCAACCTG 180
   |||||||
Db 5007 ATTACATCAACCAACCACTACTAGTTCACACACAGGAGTCAAGTCAACCTG 5066
OY 181 CAACCCACAACAGTCAAGCTAAAAACACACACACCAACCAACCCAGCAAGCC 240
   |||||||
Db 5067 CAATCCACAACATCAAGACCAACCAACCAACCAACCAACCAACCAAGCC 5126
OY 241 ACTACAAAACACGCAACCAACCAACCAACCAACCAACCAACCAATGATTTTCACTTGGAA 300
   |||||||
Db 5127 ACCCAAAAACACGCAACCAACCAACCAACCAACCAACCAACCAATGATTTTCACTTGGAA 5186
OY 301 GTGTTAACTTTGTACCTGCAACATATGCAACCAACCAACCTGCTGGCTATCTGC 360
   |||||||
Db 5187 GTGTTAACTTTGTACCTGCAACATATGCAACCAACCAACCTGCTGGCTATCTGC 5246
OY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
   |||||||
Db 5247 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5306
OY 421 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
   |||||||
Db 5307 ACCTTCAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5366
OY 481 ACCCAAGGCGCCAGACAGGCGCCAGACCAACCAACCAACCAACCAACCAACCAACCA 540
   |||||||
Db 5367 ACCCAAGGCGCCAGACAGGCGCCAGACCAACCAACCAACCAACCAACCAACCAACCA 5426
OY 541 CTGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
   |||||||
Db 5427 CTGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5486
OY 601 TCACACTCTCTCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCAGACCCA 660
   |||||||
Db 5487 TCACACTCTCTCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCACAACTCCAGACCCA 5546
OY 661 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 714
   |||||||
Db 5547 TCACAAACCTCATCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCA 5600

```

RESULT 10
AAV17553
ID AAV17553 standard; cDNA; 15223 BP.

```

XX AC AAV17553;
XX
XX 20-JUL-1998 (first entry)
XX
XX Respiratory syncytial virus antigenome.
XX
XX RSV; attenuation; vaccine; pneumonia; bronchiolitis; ss.
XX
XX Human respiratory syncytial virus D46.
XX
XX W09802530-A1.
XX
XX 22-JAN-1998.
XX
XX 15-JUL-1997; 97WO-US12269.
XX
XX 23-MAY-1997; 97US-0047634.
XX
XX 15-JUL-1996; 96US-0021773.
XX
XX 09-MAY-1997; 97US-0046141.

```

```

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Bukreyev AA, Collins PL, Juhász K, Murphy BR, Teng MN;
XX PI Whitehead SS;
XX WPI: 1998-110579/10.
XX
XX Attenuated respiratory syncytial virus vaccines - useful to protect
XX individuals against RSV infection
XX
XX Example 7, Page 188-195; 238pp; English.
XX
XX This is the 5'-3' positive sequence nucleotide sequence of
XX respiratory syncytial virus (RSV) D46. The genome is
XX negative-sense; the complete nucleotide sequence of the
XX wild-type B-1 virus has also been determined (see AAV17552).
XX A novel infectious recombinant RSV comprises a RSV genome or
XX antigenome, a major nucleocapsid (N) protein, a nucleocapsid
XX phosphoprotein (P), a large polymerase protein (L), and a RNA
XX polymerase elongation factor, where the recombinant RSV has at
XX least two attenuating mutations, one of the mutations specifying a
XX temperature-sensitive (ts) substitution at amino acid Phe521.
XX Gln831, Met1169 or Tyr1321 in the RSV polymerase gene or a ts
XX nucleotide substitution in the gene-start sequence of gene M2.
XX Also claimed are: (1) an isolated infectious RSV particle which
XX comprises a recombinant RSV (anti)genome, N, P, and L proteins, a
XX RNA polymerase elongation factor, where the (anti)genome is modified:
XX (i) to ablate or modulate expression of a SH, NS1, NS2 or G gene or
XX a cis-acting regulatory sequence; and (ii) by a termination codon
XX introduced within a selected gene, or by a change in sequence,
XX position or presence of a GS or GE transcription signal relative to
XX the selected gene; (2) an expression vector; and (3) an RSV strain
XX selected from cpts RSV 248 (ATCC VR 2450), cpts 248/404 (ATCC VR
XX 2454), cpts 248/955 (ATCC VR 2453), cpts RSV 530 (ATCC VR 2452),
XX cpts 530/1009 (ATCC VR 2451) or cpts 530/1030 (ATCC VR 2455), or
XX B-1 cp52/2B5 (ATCC VR 2542) or B-1 cp-23 (ATCC VR). The isolated
XX attenuated recombinant RSV and RSV particles are used in a vaccine
XX to stimulate the immune system of an individual to induce
XX protection against RSV. The expression vector of (2) is used for
XX the production of infectious attenuated RSV particles.
XX
XX Sequence 15223 BP; 5921 A; 2704 C; 2361 G; 4236 T; 1 other:
XX
XX Query Match 91.6%; Score 654.8; DB 19; Length 15223;
XX Best Local Similarity 94.8%; Pred. No. 2e-147;
XX Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

```

Db 5247 AAAAGATACCAAAAAAACAGGAAAGAAACACACACAGCCCAAAAAAACCA 5306
OY 421 ACCTTGACAGCAACCAAAAAAGATCTCAAAACCTCAAAACCAAGCAAGTACCC 480
Db 5307 ACCCTCAACACCAACCAAAAAAGATCTCAAAACCTCAAAACCAAGTACCC 5366
OY 481 ACCACCAAGCCCAAGAGAGCCCAACATCAACACCAACCAACCAACATCACTACA 540
Db 5367 ACCACCAAGCCCAAGAGAGCCCAACATCAACACCAACCAACCAACATCACTACA 5426
OY 541 CTGCTCACCACACACACAGCAAAATCCAAATCTCAACAGTCAAAATGGAAACCTTCAC 600
Db 5427 CTACTCTACTCCACACACAGCAAAATCCAAATCTCAACAGTCAAAATGGAAACCTTCAC 5486
OY 601 TCACACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACAAATCCGAGCAACCA 660
Db 5487 TCACACTCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCAACAAATCCGAGTACCA 5546
OY 661 TCACAAACCTCTCTCCACCAACACACACGCGAGTAGTATTAAAAAATA 714
Db 5547 TCACAAACCTCTCTCCACCAACACACACGCGAGTAGTATTAAAAAATA 5600

```

RESULT 11

AAA88743 standard; cDNA: 15223 BP.

AAA88743;

19-FEB-2001 (first entry)

Respiratory syncytial virus Da6 5'-3' positive sense sequence.

RSV; vaccine; attenuation; pneumonia; bronchiolitis; mutant; ss.

OS Chimeric - Human respiratory syncytial virus.

OS Chimeric - Bacteriophage T7.

Key Location/Qualifiers

FT FT variation

FT FT 1099

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

FT FT mutation

PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Collins PL, Murphy BR, Whitehead SS;
 XX WPI: 2000-679462/66.
 XX
 XX Infectious chimeric respiratory syncytial virus (RSV) produced from
 PT cloned nucleotide sequences, useful as a vaccine against diseases
 PT caused by the virus, such as pneumonia and bronchiolitis -
 XX
 XX Example 7; Page 262-268; 280pp; English.

The present sequence is that of the 5' to 3' positive-sense
 CC sequence of human respiratory syncytial virus (RSV); the genome
 CC itself is negative-sense. This antigenome cDNA, termed da6, was
 CC synthesized in segments by RT-PCR using synthetic oligonucleotides
 CC as primers and intracellular RSV mRNA or genome RNA isolated from
 CC purified virions as template. The antigenome includes a 5'-terminal
 CC nonviral 6 triplet contributed by the T7 promoter, 4 sequence
 CC markers (see AAA88745-47) at positions 1099 (which adds 1 nucleotide
 CC to the length), 1139, 5611 and 7559, a ribozyme and tandem T7
 CC terminators, and a single nonviral 3'-phosphorylated residue
 CC contributed to the 3' end by ribozyme cleavage. The invention
 CC provides an isolated infectious chimeric RSV comprising a major
 CC nucleocapsid (N) protein, a nucleocapsid phosphoprotein (P), a
 CC large polymerase protein (L), an RNA polymerase elongation factor,
 CC and a partial or complete RSV genome or antigenome of one RSV
 CC strain or subgroup virus combined with a heterologous gene of a
 CC different RSV strain or subgroup virus. The chimeric RSV is
 CC infectious and attenuated, preferably by introduction of selected
 CC mutations. It is useful as a vaccine against RSV, which causes
 CC diseases such as pneumonia and bronchiolitis in infants. The
 CC immune system of an individual is stimulated to induce protection
 CC against natural RSV infection, preferably in a multivalent manner
 CC to achieve protection against multiple RSV strains and/or subgroups.
 XX

Sequence 15223 BP; 5921 A; 2705 C; 2361 G; 4236 T; 0 other;

Query Match 91.6%; Score 654.8; DB 21; Length 15223;

Best Local Similarity 94.8%; Pred. No. 2e-147; Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```

OY 1 CACAAAGTCACACCTAAGCAATCATACAGATGCAACAGCCAGATCAAGACACA 60
Db 4887 CACAAAGTCACACCAACCAACATCATACAGATGCAACAGCCAGATCAAGACACA 4946
OY 61 ACCCCACATACCTCAGATCAGATCCTCAGCTTGGATGATGCTTCTGCTGAA 120
Db 4947 ACCCCACATACCTCAGATCAGATCCTCAGCTTGGATGATGCTTCTGCTGAA 5006
OY 121 ATTACATCACAACACACACATCTAGCTTCAACAACACAGAGTCAAGTCAAACTG 180
Db 5007 ATTACATCACAACACACACATCTAGCTTCAACAACACAGAGTCAAGTCAAACTG 5066
OY 181 CAACCCACACAGTCAAGACTATAAACCAACAACCAACCAACCAACCAACCAACCC 240
Db 5067 CAATCCACACAGTCAAGACTATAAACCAACAACCAACCAACCAACCAACCAACCC 5126
OY 241 ACTCAACAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
Db 5127 ACTCAACAACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5186
OY 301 GTGTTAACTTTTACCTCAGATATGAGAGCAACCAACCAACCAACCAACCAACCAAC 360
Db 5187 GTGTTAACTTTTACCTCAGATATGAGAGCAACCAACCAACCAACCAACCAACCAAC 5246
OY 361 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 420
Db 5247 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5306
OY 421 ACCTCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
Db 5307 ACCTCAAGAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5366

```

QY 481 ACCACCAAGCCGACAGAGAGCCACCATCAACACACCAAAACAAATCACTACATA 540
|||||
Db 5367 ACCACCAAGCCGACAGAGAGCCACCATCAACACACCAAAACAAATCACTACATA 5426
QY 541 CTGCTCAGCAACACACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 600
|||||
Db 5427 CTACTCAGCTCCAGACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 5486
QY 601 TCACACTCTCCGAGAGCAATCTAAGCCCTTCCAAAGTCTCCACAAATCCGACACCA 660
|||||
Db 5487 TCACACTCTCCGAGAGCAATCTAAGCCCTTCCAAAGTCTCCACAAATCCGACACCA 5546
QY 661 TCACAAACCCGATCTCCACACACACACACAGCCAGTAGTTATTTAAAAAAA 714
|||||
Db 5547 TCACAACTTCTATCTCCACACACACACACAGCCAGTAGTTATTTAAAAACATA 5600

RESULT 12
AAC88494
ID AAC88494 standard; DNA; 918 BP.
XX
AC AAC88494;
XX
DT 12-MAR-2001 (first entry)
XX
DE Human RSV G-protein gene.
XX
KW Chimeric: respiratory syncytial virus; RSV; immunize; ds.
XX
OS Unidentified.
XX
PN WO200068392-A1.
XX
PD 16-NOV-2000.
XX
PF 10-MAY-2000; 2000WO-US12582.
XX
PR 11-MAY-1999; 99US-0133536.
XX
PA (UNIT) UNIV ILLINOIS FOUND.
XX
PI Buelow DE, Korban SS, Sandhu J, Krasnyanski SF;
PI MPI; 2001-122707/13.
XX
DR Chimeric nucleic acid construct for immunizing animals and humans
PT against respiratory syncytial virus (RSV), comprises a sequence adapted
PT for expression in plants and a RSV protein or peptide coding sequence
PT
XX
PS Disclosure; Fig 10; 67pp; English.
XX
CC The present invention relates to a chimeric nucleic acid construct
CC comprising: a nucleotide sequence adapted for protein expression in
CC plants; and a respiratory syncytial virus (RSV) coding sequence
CC encoding an RSV protein or an antigenic protein or peptide of RSV.
CC The construct can be used to immunize animals and humans against
CC respiratory syncytial virus. The use of transgenic plants to
CC generate the antigen allows the production of greater amounts of
CC antigen.
XX
SQ Sequence 918 BP; 366 A; 293 C; 100 G; 159 T; 0 other;

Query Match 90.9%; Score 649.6; DB 22; Length 918;
Best Local Similarity 95.2%; Pred. No. 1.7e-146;
Matches 670; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTGCAATCATATCAAGATGCAACAGCCAGATCAAGACACA 60
|||||
Db 214 CACAAAGTCACACTGCAATCATATCAAGATGCAACAGCCAGATCAAGACACA 273
QY 61 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGCTTCCCAATCTGTCTGAA 120

Db 274 ACCCAACATACCTCACTCAGATCTCTCAGCTTGGAATCAGCTTCCCAATCTGTCTGAA 333
|||||
QY 121 ATTACATACCAACACACACACCATAGCTTCAACCAACACAGGAGTCAATCAACCTG 180
|||||
Db 334 ATTACATACCAACACACACACCATAGCTTCAACCAACACAGGAGTCAATCAACCTG 333
QY 181 CAACCAACACAGTCAAGACTTAAACACACCAACCAACCAACCAACCAACCAACCAACCC 240
|||||
Db 394 CAATTCACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 453
QY 241 ACTACAAACACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 300
|||||
Db 454 ACCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCC 513
QY 301 GTGTTAACTTTGATCCCTGAGATATGAGCAACCAATCCAACTGCTGGGTATCTGC 360
|||||
Db 514 GTGTTAACTTTGATCCCTGAGATATGAGCAACCAATCCAACTGCTGGGTATCTGC 573
QY 361 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
|||||
Db 574 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 633
QY 421 ACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
|||||
Db 634 ACCCTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 693
QY 481 ACCCAACGCCCACAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
|||||
Db 694 ACCCAACGCCCACAGAGAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 753
QY 541 CTGCTCAGCAACACACACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 600
|||||
Db 754 CTACTCAGCTCCAAACACACACAGAGAAATCCAAAAGTCACAAATGGAACCTTCCAC 813
QY 601 TCACCTCTCTCCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCACAAATCCGACACCA 660
|||||
Db 814 TCACCTCTCTCCGAGAGCAATCTAAGCCCTTCTCAAGTCTCCACAAATCCGAGTACCA 873
QY 661 TCACAAACCTCATCTCCACACACACAGCCAGTAGTTATTT 704
|||||
Db 874 TCACAACTTCTATCTCCACACACACACAGCCAGTAGTTACT 917

RESULT 13
AAS20145
ID AAS20145 standard; DNA; 897 BP.
XX
AC AAS20145;
XX
DT 09-APR-2002 (first entry)
XX
DE Respiratory syncytial virus G protein DNA.
XX
KW RSV; ds; G protein; heavily glycosylated protein; antihaemic; antiviral;
KW vaccine; gene therapy; paramyxovirus; sendai virus; PMV;
KW antiviral chemotherapeutic compound; humoral response;
KW cellular immune response; hPIV; paediatric respiratory disease;
KW globin gene transfer; sickle cell disease; beta-thalassemia;
KW human immunodeficiency virus infection; HIV.
XX
OS Human respiratory syncytial virus.
XX
FH Key 1..897 Location/Qualifiers
FH CDS 1..897 /tag= a
FH FT /product= "G protein"
XX
WO200192548-A2.
XX
PD 06-DEC-2001.
XX
PF 22-MAY-2001; 2001WO-US16610.

XX 01-JUN-2000; 2000US-208701P.
 PR (SUD-) ST JUDE CHILDREN'S RES HOSPITAL.
 XX
 PA Portner A, Takimoto T;
 XX
 DR WPI: 2002-130534/17.
 DR P-PSDB: AAU74676.
 XX
 PT Recombinant Sendai virus useful in vaccines to protect infection by
 PT paramyxoviruses, comprises exogenous nucleic acid encoding
 PT paramyxovirus protein or its antigenic fragment
 PS
 PS Disclosure: Page 47; 57pp; English.
 XX
 CC The invention relates to a recombinant Sendai virus comprising an
 CC exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its
 CC antigenic fragment. The virus may be administered in combination
 CC with an antiviral chemotherapeutic compound. Two or more viruses
 CC expressing different PMV proteins may be co-administered. Compositions
 CC comprising the virus are useful for eliciting a humoral and/or
 CC cellular immune response to a PMV in a mammal, particularly a human.
 CC Further a recombinant Sendai virus comprising an exogenous nucleic acid
 CC encoding a second PMV protein is also administered and priming and/or
 CC boosting humoral or cellular immune response comprises administering
 CC one or more of a recombinant or isolated PMV protein or its antigenic
 CC fragment, a DNA vaccine encoding the same, and a non-Sendai viral
 CC vector encoding a PMV protein. The recombinant virus is useful as an
 CC effective vaccine against hpiV or RSV (the major causes of paediatric
 CC respiratory disease) and also to express any gene of
 CC interest in target cells, providing a positive medical impact on
 CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)
 CC into stem cells effects a cure for sickle cell disease or beta-
 CC thalassemia. The recombinant virus may also prove effective in
 CC conferring immunity to human immunodeficiency virus (HIV) infection.
 CC The Sendai virus replicates at level that is high enough to
 CC induce sufficient immunity, but does not cause any harm to human
 CC recipient. The present sequence encodes a respiratory syncytial
 CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein
 CC suitable for expression by the recombinant virus of the invention.
 CC
 XX
 XX Sequence 897 BP; 359 A; 289 C; 95 G; 154 T; 0 other;
 SQ
 Query Match 90.4%; Score 646.2; DB 24; Length 897;
 Best Local Similarity 95.3%; Pred. No. 1.1e-145;
 Matches 666; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTAGCACTGCAATCATCAAGATGCAACAAAGCCAGATCAAGAACACA 60
 DB 199 CACAAAGTCACACCAACCACTGCAATCATCAAGATGCAACAAAGCCAGATCAAGAACACA 258
 QY 61 ACCCCACATACCTCAGTCTGAGATCCTGAGTGGATGAGTCTTCCATATGCTCTGAA 120
 DB 259 ACCCCACATACCTCAGTCTGAGATCCTGAGTGGATGAGTCTTCCATATGCTCTGAA 318
 QY 121 ATTACATTCACCAACCACTACTAGTTCACCAACCAAGGAGTCAAGTCAAACTGTG 180
 DB 319 ATTACATTCACCAACCACTACTAGTTCACCAACCAAGGAGTCAAGTCAAACTGTG 378
 QY 181 CAACCACACAGTCAGTCAAGCTAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 240
 DB 379 CAATCCACACAGTCAGTCAAGCTAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 438
 QY 241 ACTACAAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
 DB 439 ACCCAACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498
 QY 301 GGTGTTAACTTTGACCTGCGAGCATATGACAGCAACCAATCCAACTGCTGGCTATCTGC 360
 DB 499 GGTGTTAACTTTGACCTGCGAGCATATGACAGCAACCAATCCAACTGCTGGCTATCTGC 558
 QY 361 AAAAATAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420

DB 559 AAAAGAAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 618
 QY 421 ACCCTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
 DB 619 ACCCTCAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 678
 QY 481 ACCACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 540
 DB 679 ACCACCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 738
 QY 541 GGTGTTAACTTTGACCTGCGAGCATATGACAGCAACCAATCCAACTGCTGGCTATCTGC 600
 DB 739 GGTGTTAACTTTGACCTGCGAGCATATGACAGCAACCAATCCAACTGCTGGCTATCTGC 798
 QY 601 TCAACCTCTCCGAAAGGCAATCTAAGCCCTTCTCAAGTCTCCAGCAACATCCGAGACCA 660
 DB 799 TCAACCTCTCCGAAAGGCAATCTAAGCCCTTCTCAAGTCTCCAGCAACATCCGAGACCA 858
 QY 661 TCACAAACCTTCAATCTCCAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 699
 DB 859 TCACAAACCTTCAATCTCCAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 897

RESULT 14
 AAV38298
 ID AAV38298 standard; DNA; 696 BP.
 XX
 AC AAV38298;
 XX
 DT 26-OCT-1998 (first entry)
 XX
 DE Respiratory syncytial virus glycoprotein G gene portion.

KM RSV, glycoprotein G; pneumococcal surface protein A; PspA;
 KM Infection; Streptococcus pneumoniae; sepsis; otitis media;
 KM meningitis; bacteraemia; pneumonia; vaccine; genetic immunisation;
 KM ss.

OS Respiratory syncytial virus.

PN W09824927-A1.

PD 11-JUN-1998.

PF 04-DEC-1997; 97WO-US22847.

PR 04-DEC-1996; 96US-0759505.

PA (UYAL-) UNIV ALABAMA.

PI Briles DE, Curiel DT, McDaniel LS;

PT WPI: 1998-333343/29.

PT Plasmid containing pneumococcal epitope for expression in eukaryotic
 PT cells - useful for eliciting immunological response to pneumococcal
 PT infection or sepsis

PS Example 1; Fig 1C; 47pp; English.

PS This is a portion of the respiratory syncytial virus glycoprotein G
 PS (RSVg) gene. It has been inserted into plasmid pCDNA3 (see AAV38297)
 PS to create plasmid pCDNA3. This plasmid contains a human
 PS cytomegalovirus immediate early promoter and the RSVg gene portion
 PS such that when an in-frame fusion is made, the resultant fusion
 PS protein may be transported to, and anchored in, a mammalian cell
 PS membrane where it can be exposed to the host immune system.
 PS Insertion of pneumococcal surface protein A (PspA) coding sequence
 PS created plasmid pPSD2601. Intramuscular immunisation of BALB/c
 PS mice with pPSD2601 induced protection against an otherwise lethal
 PS challenge with a capsular type 3 pneumococcus. A claimed plasmid
 PS for expression of pneumococcal epitope DNA in eukaryotic cells

Query Match 100.0%; Score 715; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 1,1e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CACAAAGTCACACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACACA 60
    |||
Db 206 CACAAAGTCACACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACACA 265
QY 61 ACCCAACATACCTGACTGATGCTGAGTCTGAGTAATGAGTCTGCAATCTGCTGAA 120
    |||
Db 266 ACCCAACATACCTGACTGAGTCTGAGTCTGAGTAATGAGTCTGCAATCTGCTGAA 325
QY 121 ATTACATCAACAACACACACACTAGCTTACACACACAGAGAGTCAAGTCAAACTG 180
    |||
Db 326 ATTACATCAACAACACACACACTAGCTTACACACACAGAGAGTCAAGTCAAACTG 385
QY 181 CAACCCACAAAGTCAAGACTTAAACACACACAAACCCAAACACACCCAGCAAGCC 240
    |||
Db 386 CAACCCACAAAGTCAAGACTTAAACACACACAAACCCAAACACACCCAGCAAGCC 445
QY 241 ACACAAACAAACGCAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 300
    |||
Db 446 ACACAAACAAACGCAACCAACCAACCAACCAACCAATATATATTTTCACTTGAA 505
QY 301 GTGTTTAACTTTGTACCTGACAGCATATGACAGCAACATCAACTGCTGGCTATCTGC 360
    |||
Db 506 GTGTTTAACTTTGTACCTGACAGCATATGACAGCAACATCAACTGCTGGCTATCTGC 565
QY 361 AAAAGATATCAACAAACAAACCAAGAAACCAACCAACCAAGCTTCAACAAACCA 420
    |||
Db 566 AAAAGATATCAACAAACCAAGAAACCAACCAACCAAGCTTCAACAAACCA 625
QY 421 ACCTTCAAGACACCAACCAACCAACCTCAACCAACCAACCAAGCAAGTACCC 480
    |||
Db 626 ACCTTCAAGACACCAACCAACCAACCTCAACCAACCAACCAAGCAAGTACCC 685
QY 481 ACCACCAAGCCACAGAAAGCCACACATCAACACCAACCAACCAACATCAACACTAGA 540
    |||
Db 686 ACCACCAAGCCACAGAAAGCCACACATCAACACCAACCAACCAACATCAACACTAGA 745
QY 541 CTGCTACACCAACACACACAGAGAAATCCAAAACCTCAACAGTAATGAAACCTTCAC 600
    |||
Db 746 CTGCTACACCAACACACACAGAGAAATCCAAAACCTCAACAGTAATGAAACCTTCAC 805
QY 601 TCACACCTCTCCGAAAGGCAATTAAGCCCTTCAAGTCTCCAGCAATCCGAGCAACCA 660
    |||
Db 806 TCACACCTCTCCGAAAGGCAATTAAGCCCTTCAAGTCTCCAGCAATCCGAGCAACCA 865
QY 661 TCACACCTCTCATCTCCACCCAAACACACAGCCAGTAGTTATTTAAAAA 715*
    |||
Db 866 TCACACCTCTCATCTCCACCCAAACACACAGCCAGTAGTTATTTAAAAA 920
```

RESULT 2

US-08-838-189D-7
; Sequence 7, Application US/08838189D
; Patent No. 5938169

GENERAL INFORMATION:
; APPLICANT: KLEIN, Michel H
; APPLICANT: DU, Run-Pan
; TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 6th Floor, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7

COMPUTER READABLE FORM:

```
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/838,189D
;; FILING DATE: 16-APR-1997
;; CLASSIFICATION: 435
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: US 08/001,554
;; FILING DATE: 06-JAN-1993
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: GB 9200117.1
;; FILING DATE: 06-JAN-1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: STEWART, Michael I
;; REGISTRATION NUMBER: 24,973
;; REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (416) 595-1155
;; TELEFAX: (416) 595-1163
;; INFORMATION FOR SEQ ID NO: 7:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 920 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
US-08-838-189D-7
```

Query Match 100.0%; Score 715; DB 2; Length 920;
Best Local Similarity 100.0%; Pred. No. 1,1e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CACAAAGTCACACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACACA 60
    |||
Db 206 CACAAAGTCACACTTAACACTGCAATCATACAGATGCAACAGCCAGATCAAGAAACACA 265
QY 61 ACCCAACATACCTGACTGAGTCTGAGTCTGAGTAATGAGTCTGCAATCTGCTGAA 120
    |||
Db 266 ACCCAACATACCTGACTGAGTCTGAGTCTGAGTAATGAGTCTGCAATCTGCTGAA 325
QY 121 ATTACATCAACAACACACACACTAGCTTACACACACAGAGAGTCAAGTCAAACTG 180
    |||
Db 326 ATTACATCAACAACACACACACTAGCTTACACACACAGAGAGTCAAGTCAAACTG 385
QY 181 CAACCCACAAAGTCAAGACTTAAACACACAAACCAACCAACCAACCAAGCAAGCC 240
    |||
Db 386 CAACCCACAAAGTCAAGACTTAAACACACAAACCAACCAACCAACCAAGCAAGCC 445
QY 241 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 300
    |||
Db 446 ACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAATATGATTTTCACTTGAA 505
QY 301 GTGTTTAACTTTGTACCTGACAGCATATGACAGCAACATCAACTGCTGGCTATCTGC 360
    |||
Db 506 GTGTTTAACTTTGTACCTGACAGCATATGACAGCAACATCAACTGCTGGCTATCTGC 565
QY 361 AAAAGATATCAACAAACAAACCAAGAAACCAACCAACCAAGCTTCAACAAACCA 420
    |||
Db 566 AAAAGATATCAACAAACCAAGAAACCAACCAACCAAGCTTCAACAAACCA 625
QY 421 ACCTTCAAGACACCAACCAACCAACCTCAACCAACCAACCAAGCAAGTACCC 480
    |||
Db 626 ACCTTCAAGACACCAACCAACCAACCTCAACCAACCAACCAAGCAAGTACCC 685
QY 481 ACCACCAAGCCACAGAAAGCCACATCAACACACCAACCAACCAACCAACCAAGTACA 540
    |||
Db 686 ACCACCAAGCCACAGAAAGCCACATCAACACACCAACCAACCAACCAACCAAGTACA 745
QY 541 CTGCTACACCAACACACACAGAGAAATCCAAAACCTCAACAGTCAAAATGAAACCTTCAC 600
    |||
Db 746 CTGCTACACCAACACACACAGAGAAATCCAAAACCTCAACAGTCAAAATGAAACCTTCAC 805
```

Accession	Sequence	Position
Oy	TCACCTCTCTCCGAAGCATCTTAAGCCCTTTCAATCTCCACAAACATCCGAGACCCA	66
Db	TCAACCTCTCTCGAAGCATCTTAAGCCCTTTCAATCTCTCCAAACATCTCCGAGACCCA	865
Oy	TCACAACTTCATCTCCACCCACACACACAGCCAGTAGTTATTTAAAAA	715
Db	TCACAACTTCATCTCCACCCACACACACAGCCAGTAGTTATTTAAAAA	920

RESULT 3

```

US-08-852-344D-7
: Sequence 7, Application US/08852344D
: Patent No. 6017539
:
GENERAL INFORMATION:
: APPLICANT: KLEIN, Michel H
: APPLICANT: DU, Run-Pan
: APPLICANT: EMASYSHYN, Mary E
: TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
: TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
: TITLE OF INVENTION: SYNCYTIAL VIRUS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6th Floor, 330 University Avenue
: City: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
:
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/852,344D
: FILING DATE: 07-MAY-1997
: CLASSIFICATION: 424
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,639
: FILING DATE: 14-NOV-1994
:
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9200117.1
: FILING DATE: 06-JAN-1992
:
ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-668 MTS:jb
:
TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
:
INFORMATION FOR SEO ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 920 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: Single
: TOPOLOGY: Linear
:
US-08-852-344D-7

```

Query Match	100.0%;	Score 715;	DB 3;	Length 920;
Best Local Similarity	100.0%;	Pred. No. 1.1e-185;		
Matches 715; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	C A A A A T C C A C A T A A C A C T G C A A T A T C A A G A T G C A A C C G C A G A T C A A G A A C A G A	60
Db	206	C A C A A A T C A C A C T A A C A A C T G C A T A T A C A A G A T C C A A C C A C C A G A T C A A G A A C A	265
QY	61	A C C C A A C A T A C C T C A C T C A G A T C C T C A C A T T G G A A T C A G T T C T C C A A T C G T G A A	120
Db	266	A C C C A A C A T A T C C A C A T C A G A T C C T C A G C T T G G A T C A G C T T C C A A T C T G T C G A A	325
QY	121	A T T A C A T C A C A A A C C A C C A C A T A C T A G C T T C A C A A C A C C A G A G T C A A G T C A A A C C T G	180

Db	326	ATTACATCCAAACACACACCATCTACTGCTTCAACCAACACACAGAGTCAAGTCAAAACCTG	385
OY	181	CAACCCACACACAGTCATAGACTTAAAAACACACAACAACAACCCAAACACACACCCAGCAAGCCC	240
Db	386	CAACCCACACACAGTCATAGACTTAAAAACACAACAACAACCCAAACACCAACCCAGCAAGCCC	445
OY	241	ACTACAAAAACAAGCCGCAAAACCAAAACCCAAACCAACCAACCCAAATATATGATTTTCCTTGAA	300
Db	446	ACTACAAAAACAAGCCGCAAAACCAAAACCCAAACCAACCAACCAATATATGATTTTCCTTGAA	505
OY	301	GTGATTAACTTTGTACCCCTGAGCATATGACAGCAACCAATCAACCTCTGGCTATCTGC	360
Db	506	GTGATTAACTTTGTATACCTCTGAGCATATGACAGCAACCAATCAACCTCTGGCTATCTGC	565
OY	361	AAAAGAATATCCAAACAAAAAACCCAGAAAGAAAAACCAACCAACCAAGCTTACAAAAAACCA	420
Db	566	AAAAGAATATCCAAACAAAAAACCCAGAAAGAAAAACCAACCAACCAAGCTTACAAAAAACCA	625
OY	421	ACCTTCAAGACACACAAAAAAGATCTCAAACTTAAACCTTAACCAAGGAATACCC	480
Db	626	ACCTTCAAGACACACAAAAAAGATCTCAAACTTCAAAACCTTAACCAAGGAATACCC	685
OY	481	ACCCACCAAGCCCAACGAAGAGCCCAACCATCAACACCCACCAAAACCAAACTATCACATCTACA	540
Db	686	ACCCACCAAGCCCAACGAAGAGCCCAACCATCAACACCCACCAAAACCAAACTATCACATCTACA	745
OY	541	CTGCTCACCAACACACACACACAGAAATCTCAAAATCTCAACAGTCAAAATGGAATCTTCAC	600
Db	746	CTGCTCACCAACCAACACACACAGAAATCTCAAAATCTCAACAGTCAAAATGGAATCTTCAC	805
OY	601	TCCAACTCTCCGGAAGGCAATCTTAAGCCCTTCTCAAGTCTTCACACACATCTCGAGCACCCA	660
Db	806	TCCAACTCTCTCGGAAGGCAATCTTAAGCCCTTCTCAAGTCTTCACACACATCTCGAGCACCCA	865
OY	661	TCCAAACCCCTCATCTCCACACCCCAACCAACACAGCCAGTAGTATTATTAACAAAAAAA	715
Db	866	TCCAAACCCCTCATCTCCACACCCCAACCAACACAGCCAGTAGTATTATTAACAAAAAAA	920

RESULT 4

US-08-344-639E-7
Sequence 7, Application US/08344639E
Patent No. 6033668
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewaszshyn, Mary E
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McHurney
STREET: 330 University Avenue, 6th Floor
City: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/344,639E
FILING DATE: 14-NOV-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1

FILING DATE: 06-JAN-1992
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-391 MIS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-344-639E-7

Query Match 100.0%; Score 715; DB 3; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.1e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTACACTGCATCAACAGATGCAACAGCAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTACACTGCATCAACAGATGCAACAGCAGATCAAGAACACA 265
QY 61 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCCATCTGTGAA 120
DB 266 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCCATCTGTGAA 325
QY 121 ATTACATCAACACACACACACTAGTACTGTTCAACACACACAGAGTCAAGTCAACCTG 180
DB 326 ATTACATCAACACACACACACTAGTACTGTTCAACACACACAGAGTCAAGTCAACCTG 385
QY 181 CAACCCACACAGTCAGATCAACACACACACACACACACACACACACACACACAC 240
DB 386 CAACCCACACAGTCAGATCAACACACACACACACACACACACACACACACACAC 445
QY 241 ACTACAAACACAGCCCAACACACACACACACACACACACACACACACACACAC 300
DB 446 ACTACAAACACAGCCCAACACACACACACACACACACACACACACACACACAC 505
QY 301 GTGTTTAACTTTGATACCTCAGCATATGACGACACACATCCAACTGCTGCTATCTGC 360
DB 506 GTGTTTAACTTTGATACCTCAGCATATGACGACACACATCCAACTGCTGCTATCTGC 565
QY 361 AAAAGATACCAACAAACAAACAGAAACAAACAAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGATACCAACAAACAAACAGAAACAAACAAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCTTCAAGACAAACAAACAAACAGTCTCAAACTCAAAACCACTAAACCAAGAGTACCC 480
DB 626 ACCTTCAAGACAAACAAACAAACAGTCTCAAACTCAAAACCACTAAACCAAGAGTACCC 685
QY 481 ACCACCAAGCCACAGAGAGCCACACATACACACACCAACCAACCAACATACACACTACA 540
DB 686 ACCACCAAGCCACAGAGAGCCACACATACACACACCAACCAACCAACATACACACTACA 745
QY 541 CTGCTACCAACACACAGAGAAATCCAAACTCAACAGTCAAGTGAAGAACTCTCCAC 600
DB 746 CTGCTACCAACACACAGAGAAATCCAAACTCAACAGTCAAGTGAAGAACTCTCCAC 805
QY 601 TCAACCTCTCTCCAGAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCA 660
DB 806 TCAACCTCTCTCCAGAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCA 865
QY 661 TCACAAACCTCATCTCCACCAACACACACAGCAGTAGTTATTTAAAAA 715
DB 866 TCACAAACCTCATCTCCACCAACACACACAGCAGTAGTTATTTAAAAA 920

RESULT 5

US-08-467-969A-7
Sequence 7, Application US/08467969A
Patent No. 6168786
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasysbyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,969A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-475 MIS:bn
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-969A-7

Query Match 100.0%; Score 715; DB 4; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.1e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCACACTACACTGCAATCTACAGATGCAACAGCAGATCAAGAACACA 60
DB 206 CACAAAGTCACACTACACTGCAATCTACAGATGCAACAGCAGATCAAGAACACA 265
QY 61 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCCATCTGTGAA 120
DB 266 ACCCAACATACCTCCTCAGATGCTCAGCTTGAGATGCTTCCATCTGTGAA 325
QY 121 ATTACATCAACACACACACACTAGTACTGTTCAACACACACAGAGTCAAGTCAACCTG 180
DB 326 ATTACATCAACACACACACACTAGTACTGTTCAACACACACAGAGTCAAGTCAACCTG 385
QY 181 CAACCCACACAGTCAGATCAACACACACACACACACACACACACACACACACAC 240
DB 386 CAACCCACACAGTCAGATCAACACACACACACACACACACACACACACACACAC 445
QY 241 ACTACAAACACAGCCCAACACACACACACACACACACACACACATATGATTTTCACTGGA 300
DB 446 ACTACAAACACAGCCCAACACACACACACACACACACACACATATGATTTTCACTGGA 505

```

QY 301 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAATCCAACTGGCTGGCTATCTGC 360
    |||||||
Db 506 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAATCCAACTGGCTGGCTATCTGC 565
QY 361 AAAAGATACCAACCAACCAAGGAAAGAAACACACACACCTACCAACCAACCA 420
    |||||||
Db 566 AAAAGATACCAACCAACCAAGGAAAGAAACACACACACCTACCAACCAACCA 625
QY 421 ACCTTCAAGACCAACCAACCAAGGAAAGAAACACACACACCTACCAACCAACCA 480
    |||||||
Db 626 ACCTTCAAGACCAACCAACCAAGGAAAGAAACACACACACCTACCAACCAACCA 685
QY 481 ACCACCAAGCCACAGAGAGCCACATCAACACACCAACCAACCAACCAACCACTACA 540
    |||||||
Db 686 ACCACCAAGCCACAGAGAGCCACATCAACACACCAACCAACCAACCAACCACTACA 745
QY 541 CTGCTCACCACACACACCAAGGAAATCCAAACTCACAAGTCAAAATGGAACCTTCAC 600
    |||||||
Db 746 CTGCTCACCACACACCAAGGAAATCCAAACTCACAAGTCAAAATGGAACCTTCAC 805
QY 601 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAATCTCCACACATCCGAGACCCA 660
    |||||||
Db 806 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 865
QY 661 TCACCAACCTCATCTCCACCAACCAACGCGCAGTAGTATTAAAAA 715
    |||||||
Db 866 TCACCAACCTCATCTCCACCAACCAACGCGCAGTAGTATTAAAAA 920

```

RESULT 6
US-08-467-961A-7
Sequence 7, Application US/08467961A
Patent No. 6171783

```

GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasashyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Slim & McBurney
STREET: 330 University Avenue, 6TH Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467, 961A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-476 MIS:bn
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 7:

```

```

SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-961A-7

```

Query Match 100.0%; Score 715; DB 4; Length 920;
Best Local Similarity 100.0%; Pred. No. 1,1e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CACAAAGTCACATCAACATCGCATATCATCAAGATGCAACGACGATCAAGACACA 60
    |||||||
Db 206 CACAAAGTCACATCAACATCGCATATCATCAAGATGCAACGACGATCAAGACACA 265
QY 61 ACCCAACATACCTACTCAGATCTCAGCTTGGATTCAGTTCCTCAATCTGTGGA 120
    |||||||
Db 266 ACCCAACATACCTACTCAGATCTCAGCTTGGATTCAGTTCCTCAATCTGTGGA 325
QY 121 ATTACATCAACACACACACCATCTAGCTTCAACAACACACGAGTCAAGTCAAACTG 180
    |||||||
Db 326 ATTACATCAACACACACCATCTAGCTTCAACAACACACGAGTCAAGTCAAACTG 385
QY 181 CAACCCACACAGTCACAGATCAAAAAACAAACAAACCCAAACACACACACCAAGCCC 240
    |||||||
Db 386 CAACCCACACAGTCACAGATCAAAAAACAAACAAACCCAAACACACACACCAAGCCC 445
QY 241 ACTACAAACACAGCCCAACCAACCAACCAACCAACCAACCAATTAATTTCTTGAA 300
    |||||||
Db 446 ACTACAAACACAGCCCAACCAACCAACCAACCAACCAACCAATTAATTTCTTGAA 505
QY 301 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAATCCAACTGGCTGGCTATCTGC 360
    |||||||
Db 506 GTGTTAACTTTGTACCCCTGCAGCATATGCAGCAATCCAACTGGCTGGCTATCTGC 565
QY 361 AAAAGATACCAACCAACCAAGGAAAGAAACACACACACCTACCAACCAACCA 420
    |||||||
Db 566 AAAAGATACCAACCAACCAAGGAAAGAAACACACACACCTACCAACCAACCA 625
QY 421 ACCTTCAAGACCAACCAACCAAGGAAAGTTCAAAGTCAACCACTCAACCAAGGAAAGTACCC 480
    |||||||
Db 626 ACCTTCAAGACCAACCAACCAAGGAAAGTTCAAAGTCAACCACTCAACCAAGGAAAGTACCC 685
QY 481 ACCACCAAGCCACAGAGAGCCACATCAACACACCAACCAACCAACCAACCACTACA 540
    |||||||
Db 686 ACCACCAAGCCACAGAGAGCCACATCAACACACCAACCAACCAACCAACCACTACA 745
QY 541 CTGCTCACCACACACCAAGGAAATCCAAACTCACAAGTCAAAATGGAACCTTCAC 600
    |||||||
Db 746 CTGCTCACCACACCAAGGAAATCCAAACTCACAAGTCAAAATGGAACCTTCAC 805
QY 601 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 660
    |||||||
Db 806 TCACCTCTCTCGAAGGCAATCTAGCCCTTCTCAAGTCTCCACACATCCGAGACCCA 865
QY 661 TCACCAACCTCATCTCCACCAACCAACGCGCAGTAGTATTAAAAA-AAAA 715
    |||||||
Db 866 TCACCAACCTCATCTCCACCAACCAACGCGCAGTAGTATTAAAAA 920

```

RESULT 7
US-08-001-554A-7
Sequence 7, Application US/08001554A
Patent No. 6225091
GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasashyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESS: Slim & McBurney

STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/001,554A
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-286
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-001-554A-7

Query Match 100.0%; Score 715; DB 4; Length 920;
Best Local Similarity 100.0%; Pred. No. 1.1e-185;
Matches 715; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCACAAAGCCCATCAAGAACCA 60
DB 206 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCACAAAGCCCATCAAGAACCA 265
QY 61 ACCCCAACTACTCTCAGATCCTCAGTCTGGAATCAGCTTCCATCTGCTGAA 120
DB 266 ACCCCAACTACTCTCAGATCCTCAGTCTGGAATCAGCTTCCATCTGCTGAA 325
QY 121 ATTACATCAACAAACCAACCATCTAGTCTTCAACAAACCAAGAGTCAAGTCAAAACCTG 180
DB 326 ATTACATCAACAAACCAACCATCTAGTCTTCAACAAACCAAGAGTCAAGTCAAAACCTG 385
QY 181 CAACCCAAAGTCAAGACTAAACCAACAAACCAACCAACCAACCAACCAACCAACCAACCA 240
DB 386 CAACCCAAAGTCAAGACTAAACCAACAAACCAACCAACCAACCAACCAACCAACCAACCA 445
QY 241 ACTACAAACCAAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
DB 446 ACTACAAACCAAGCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 505
QY 301 GTGTTTAACCTTTGTACCTGACAGCATATGACAGCAATCCAACTGCTGGGCTATCTGC 360
DB 506 GTGTTTAACCTTTGTACCTGACAGCATATGACAGCAATCCAACTGCTGGGCTATCTGC 565
QY 361 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 420
DB 566 AAAAGATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 625
QY 421 ACCCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 480
DB 626 ACCCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 685
QY 481 ACCACCAAGCCACAGAAAGCCAAACCATATACAGCAACCAACCAACCAACCAACCAACCA 540
DB 686 ACCACCAAGCCACAGAAAGCCAAACCATATACAGCAACCAACCAACCAACCAACCAACCA 745
QY 541 CTGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600

DB 746 CTGCTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 805
QY 601 TCACACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCAGACCA 660
DB 806 TCACACCTCTCCGAGGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCAGACCA 865
QY 661 TCACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 715
DB 866 TCACACCTCTCTCCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 920
RESULT 8
US-08-467-963C-28
Sequence 28, Application US/08467963C
Patent No. 5968776
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,963C
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/838,189
FILING DATE: 16-Apr-1997
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-467-963C-28
Query Match 97.3%; Score 696; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 1.6e-180;
Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCACAAAGCCCATCAAGAACCA 60
DB 199 CACAAAGTACACTAACAACCTGCAATCATCAAGATGCACAAAGCCCATCAAGAACCA 258
QY 61 ACCCCAACTACTCTCAGATCCTCAGTCTGGAATCAGCTTCCATCTGCTGAA 120
DB 259 ACCCCAACTACTCTCAGATCCTCAGTCTGGAATCAGCTTCCATCTGCTGAA 318

QY 121 ATTACATCAACAAACCAACATAGTCTGACAGACAGAGTCAAGTCAAACTG 180
|||||
Db 319 ATTACATCAACAAACCAACATAGTCTGACAGACAGAGTCAAGTCAAACTG 378
QY 181 CAACCCACAGAGTCAAGCTAAAAACACACACACCAACCAACCAAGAGCC 240
|||||
Db 379 CAACCCACAGAGTCAAGCTAAAAACACACACCAACCAACCAAGAGCC 438
QY 241 ACTACAAACAGAGCCAAACCAACACCAACCAACCAATATGATTTTCACTTCAA 300
|||||
Db 439 ACTACAAACAGAGCCAAACCAACACCAACCAACCAATATGATTTTCACTTCAA 498
QY 301 GTGTTAACTTGTACCTGACGATATGACAGCAACATGCAACCTGCTGGCTATGCG 360
|||||
Db 499 GTGTTAACTTGTACCTGACGATATGACAGCAACATGCAACCTGCTGGCTATGCG 558
QY 361 AAAAGATATCAACAAACCAACCAAGAAAGAAACCAACCAAGCTTACAAACCA 420
|||||
Db 559 AAAAGATATCAACAAACCAAGAAAGAAAGAAACCAACCAAGCTTACAAACCA 618
QY 421 ACCTTCAAGACAAACCAACCAAGATCTCAACCTCAACCAAGCAAGCAAGTACCC 480
|||||
Db 619 ACCTTCAAGACAAACCAACCAAGATCTCAACCTCAACCAAGCAAGCAAGTACCC 678
QY 481 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAAGCAAGTACCA 540
|||||
Db 679 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAAGCAAGTACCA 738
QY 541 CTGCTACCAACCAACCAAGAGAAATCCAAATCTCAAGTCAATGAAACCTTCCAG 600
|||||
Db 739 CTGCTACCAACCAACCAAGAGAAATCCAAATCTCAAGTCAATGAAACCTTCCAG 798
QY 601 TCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCACAACATCCGAGACCCA 660
|||||
Db 799 TCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCACAACATCCGAGACCCA 858
QY 661 TCACAACCTCTATCTCCACCAACCAACCAAGCCAG 696
|||||
Db 859 TCACAACCTCTATCTCCACCAACCAACCAAGCCAG 894

RESULT 9

US-08-838-189D-28
Sequence 28, Application US/08838189D
Patent No. 5998169
GENERAL INFORMATION:
APPLICANT: KLEIN, Michel H
APPLICANT: DU, Run-Pan
APPLICANT: EMASYSYN, Mary E
TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 6th Floor, 330 University Avenue
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838,189D
FILING DATE: 16-Apr-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-Jan-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-Jan-1992
ATTORNEY/AGENT INFORMATION:
NAME: STEWART, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-687 MIS:jb
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: Linear
US-08-838-189D-28

Query Match 97.3%; Score 696; DB 2; Length 894;
Best Local Similarity 100.0%; Pred. No. 1,6e-180;

Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTCAACACTAGCAATGATCAAGATGCAACCAAGCAAGATCAAGTACACA 60
|||||
Db 199 CACAAAGTCAACACTAGCAATGATCAAGATGCAACCAAGCAAGATCAAGTACACA 258
QY 61 ACCCAACATACCTGATCGATGAGATCTGACGTTGGATTCAGCTTCTGCAATCTGCTGAA 120
|||||
Db 259 ACCCAACATACCTGATCGATGAGATCTGACGTTGGATTCAGCTTCTGCAATCTGCTGAA 318
QY 121 ATTACATCAACAAACCAACCAACATAGTCTTCAACCAACCAAGAGTCAACCTCAACCTG 180
|||||
Db 319 ATTACATCAACAAACCAACCAACATAGTCTTCAACCAACCAAGAGTCAACCTCAACCTG 378
QY 181 CAACCCACAGAGTCAAGCTAAAAACACACCAACCAACCAAGCAAGCTTACAAACCTG 240
|||||
Db 379 CAACCCACAGAGTCAAGCTAAAAACACACCAACCAACCAAGCAAGCTTACAAACCTG 438
QY 241 ACTACAAACAGAGCCAAACCAACCAACCAACCAACCAATATGATTTTCACTTCAA 300
|||||
Db 439 ACTACAAACAGAGCCAAACCAACCAACCAACCAACCAATATGATTTTCACTTCAA 498
QY 301 GTGTTAACTTGTACCTGACGATATGACAGCAACATGCAACCTGCTGGCTATGCG 360
|||||
Db 499 GTGTTAACTTGTACCTGACGATATGACAGCAACATGCAACCTGCTGGCTATGCG 558
QY 361 AAAAGATATCAACAAACCAACCAAGAAAGAAACCAACCAAGCTTACAAACCA 420
|||||
Db 559 AAAAGATATCAACAAACCAAGAAAGAAAGAAACCAACCAAGCTTACAAACCA 618
QY 421 ACCTTCAAGACAAACCAACCAAGATCTCAAACTCAACCAAGCAAGTAAACCAAGAGTACCC 480
|||||
Db 619 ACCTTCAAGACAAACCAACCAAGATCTCAAACTCAACCAAGCAAGTAAACCAAGAGTACCC 678
QY 481 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAAGCAAGTACCA 540
|||||
Db 679 ACCACCAAGCCCAAGAGAGCAACATCAACACCAACCAACCAAGCAAGTACCA 738
QY 541 CTGCTACCAACCAACCAAGAGAAATCCAAATCTCAAGTCAATGAAACCTTCCAG 600
|||||
Db 739 CTGCTACCAACCAACCAAGAGAAATCCAAATCTCAAGTCAATGAAACCTTCCAG 798
QY 601 TCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCACAACATCCGAGTACCCA 660
|||||
Db 799 TCAACCTCTCGAAGGCAATCTAAGCCCTTCTCAAGTCTCACAACATCCGAGTACCCA 858
QY 661 TCACAACCTCTATCTCCACCAACCAACCAAGCCAG 696
|||||
Db 859 TCACAACCTCTATCTCCACCAACCAACCAAGCCAG 894

RESULT 10

US-08-852-344D-28

```

: Sequence 28, Application US/08852344D
: Patent No. 6017539
: GENERAL INFORMATION:
: APPLICANT: KLEIN, Michel H
: APPLICANT: DU, Run-Pan
: APPLICANT: EWASYSHYN, Mary E
: TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
: TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
: TITLE OF INVENTION: SYNCYTIAL VIRUS
: NUMBER OF SEQUENCES: 37
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 6th Floor, 330 University Avenue
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/852,344D
: FILING DATE: 07-MAY-1997
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/344,639
: FILING DATE: 14-NOV-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 9200117.1
: FILING DATE: 06-JAN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: STEWART, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-688 MIS-Jb
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: INFORMATION FOR SEQ ID NO: 28:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 894 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: US-08-852-344D-28
:
: Query Match 97.3%; Score 696; DB 3; Length 894;
: Best Local Similarity 100.0%; Pred. No. 1.6e-180;
: Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 499 GTGTTTAACCTTTGACCTGACGATATGAGCAACAATCCAACTGCTGGCTATCTGCG 558
QY 361 AAAGAAATACCAACCAAAAAACCGAAAGAAACCGACCAAGCTTACAAAAACCA 420
Db 559 AAAAGAAATACCAACCAAAAAACCGAAAGAAACCGACCAAGCTTACAAAAACCA 618
QY 421 ACCTTCAAGCAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCAAGGAATACC 480
Db 619 ACCTTCAAGCAACCAAAAAAGATCTCAAACTCAAAACCACTAAACCAAGGAATACC 678
QY 481 ACCACCAAGCCACAGAGAGCCACCATATACACCAACCAACCAACCAATCACTACCA 540
Db 679 ACCACCAAGCCACAGAGAGCCACCATATACACCAACCAACCAACCAATCACTACCA 738
QY 541 CTGCTACCAACCAACCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 600
Db 739 CTGCTACCAACCAACCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 798
QY 601 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCACTTCACACATCCGAGCACCA 660
Db 799 TCAACCTCTCCGAAGGCAATCTAAGCCCTTCTCACTTCACACATCCGAGCACCA 858
QY 661 TCACAAACCTCATCTCCACCCCAACCAACCAAGCCAG 696
Db 859 TCACAAACCTCATCTCCACCCCAACCAACCAAGCCAG 894

```

```

: RESULT 11
: US-08-344-639E-28
: Sequence 28, Application US/08344639E
: Patent No. 603568
: GENERAL INFORMATION:
: APPLICANT: Klein, Michel H
: APPLICANT: Du, Run-Pan
: APPLICANT: EWASYSHYN, Mary E
: TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS
: TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
: TITLE OF INVENTION: AND RESPIRATORY SYNCYTIAL VIRUS
: NUMBER OF SEQUENCES: 38
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Sim & McBurney
: STREET: 330 University Avenue, 6th Floor
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5G 1R7
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/344,639E
: FILING DATE: 14-NOV-1994
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/001,554
: FILING DATE: 06-JAN-1992
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Stewart, Michael I
: REGISTRATION NUMBER: 24,973
: REFERENCE/DOCKET NUMBER: 1038-391 MIS
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 595-1155
: TELEFAX: (416) 595-1163
: TELEX: 065-24567 SIMBAS
: INFORMATION FOR SEQ ID NO: 28:
: SEQUENCE CHARACTERISTICS:

```

LENGTH: 894 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-344-639E-28

Query Match 97.3%; Score 696; DB 3; Length 894;
 Best Local Similarity 100.0%; Pred. No. 1.6e-180;
 Matches 696; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAAAGTACACTACAACTGCAATCATACAGATGCAAGCAGCATCAAGAACACA 60
 DB 199 CACAAAGTACACTACAACTGCAATCATACAGATGCAAGCAGCATCAAGAACACA 258
 QY 61 ACCCAACATACCTCAGATGCTGAGTGGAAATCAGCTTCTCCAAATCTCTGAA 120
 DB 259 ACCCAACATACCTCAGATGCTGAGTGGAAATCAGCTTCTCCAAATCTCTGAA 318
 QY 121 ATTACATCAAAACACCCAGCATCTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 180
 DB 319 ATTACATCAAAACACCCAGCATCTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 378
 QY 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240
 DB 379 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 438
 QY 241 ACTACAAACACAGGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
 DB 439 ACTACAAACACAGGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 498
 QY 301 GTGTTAACTTTGTACCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 360
 DB 499 GTGTTAACTTTGTACCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 558
 QY 361 AAAAGAAATACCAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 420
 DB 559 AAAAGAAATACCAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 618
 QY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
 DB 619 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 678
 QY 481 ACCCAACAGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540
 DB 679 ACCCAACAGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 738
 QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 600
 DB 739 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 798
 QY 601 TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCAGACATCCGAGCAGCA 660
 DB 799 TCAACCTCTCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCAGACATCCGAGCAGCA 858
 QY 661 TCACACCTCTATCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 696
 DB 859 TCACACCTCTATCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 894

RESULT 12

US-08-801-898A-23

Sequence 23; Application US/08801898A

Patent No. 5998602

GENERAL INFORMATION:

APPLICANT: Torrence, Paul F.

APPLICANT: Silverman, Robert H.

APPLICANT: Cirino, Nick M.

APPLICANT: Li, Guliying

APPLICANT: Xiao, Wei

TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE

TITLE OF INVENTION: OLIGONUCLEOTIDES EFFECTIVE TO TREAT RSV INFECTIONS

NUMBER OF SPOUNCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: NY
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/801,898A
 FILING DATE: 18-FEB-1997
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Poissant, Brian M.
 REGISTRATION NUMBER: 28,462
 REFERENCE/DOCKET NUMBER: 8656-009
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-790-9090
 TELEFAX: 212-869-9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 23:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15222 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 FEATURE:
 NAME/KEY: RSV-A2
 LOCATION: 1..15222
 OTHER INFORMATION:
 US-08-801-898A-23

Query Match 91.6%; Score 654.8; DB 2; Length 15222;
 Best Local Similarity 94.8%; Pred. No. 8.5e-169;
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAAGTCACTAACAATGCAATCATACAGATGCAAGCAGCATCAAGAACACA 60
 DB 4886 CACAAAGTCACTAACAATGCAATCATACAGATGCAAGCAGCATCAAGAACACA 4945
 QY 61 ACCCAACATACCTCAGATGCTGAGTGGAAATCAGCTTCTCCAAATCTCTGAA 120
 DB 4946 ACCCAACATACCTCAGATGCTGAGTGGAAATCAGCTTCTCCAAATCTCTGAA 5005
 QY 121 ATTACATCAAAACACCCAGCATCTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 180
 DB 5006 ATTACATCAAAACACCCAGCATCTAGCTTCAACAAACAGGAGTCAAGTCAAACTG 5065
 QY 181 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 240
 DB 5066 CAACCCCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 5125
 QY 241 ACTACAAACACAGGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 300
 DB 5126 ACTACAAACACAGGCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 5185
 QY 301 GTGTTAACTTTGTACCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 360
 DB 5186 GTGTTAACTTTGTACCTGAGATATGACAAATCCACCTGCTGGCTATCTGC 5245
 QY 361 AAAAGAAATACCAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 420
 DB 5246 AAAAGAAATACCAAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 5305
 QY 421 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
 DB 5306 ACCTTCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5365
 QY 481 ACCCAACAGCCCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540

|||||
Db 5366 ACCACCAAGCCACAGAGAGCCATCAACACACCAAAACATCATATACATCA 5425
OY 541 CTGCTACCAACACACACAGAGAAATCCAAATCTCAAGTCAATGGAACCTTCAC 600
Db 5426 CTACTACCTCCACACACAGAGAAATCCAGAACTCAAGTCAATGGAACCTTCAC 5485
OY 601 TCACCTCTCCGAGAGCAATCTAGCCCTCTCAAGTCTCCACAACTCCAGACCA 660
Db 5486 TCACCTCTCCGAGAGCAATCCAGCCCTTCTCAAGTCTCTACACATCCGAGACCA 5545
OY 661 TCACACCTCTATCTCCACCCACACACAGCCAGTAGTATTAAAAAAA 714
Db 5546 TCACACCTCTATCTCCACCCACACACAGCCAGTAGTATTAAAAACATA 5599
RESULT 13
US-08-962-690-12/C
; Sequence 12, Application US/08962690
; Patent No. 6214805
; GENERAL INFORMATION:
; APPLICANT: Torrence, Paul F.
; APPLICANT: Silverman, Robert H.
; APPLICANT: Cirino, Nick M.
; APPLICANT: Li, Gulying
; APPLICANT: Xiao, Wei
; APPLICANT: Player, Mark R.
; TITLE OF INVENTION: RNASE L ACTIVATORS AND ANTISENSE OLIGONUCLEOTIDES
; TITLE OF INVENTION: EFFECTIVE TO TREAT RSV INFECTIONS
; FILE REFERENCE: 8656-019
; CURRENT APPLICATION NUMBER: US/08/962,690
; EARLIER FILING DATE: 1997-11-03
; EARLIER APPLICATION NUMBER: 08/801,896
; EARLIER FILING DATE: 1997-02-14
; EARLIER APPLICATION NUMBER: 60/011,725
; EARLIER FILING DATE: 1996-02-15
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 12
; LENGTH: 15222
; TYPE: DNA
; ORGANISM: respiratory syncytial virus
US-08-962-690-12
Query Match 91.6%; Score 654.8; DB 4; Length 15222;
Best Local Similarity 94.8%; Pred. No. 8.5e-169;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

|||||
Db 9977 AAAAGATACCAAAACAAAAACAGAGAAAGAACCACTACCAAGCCACAAAAACCA 9918
OY 421 ACCTTCAAGCAACCAAAAAAGATCTCAACCTCAAAACCACTAAACCAAGGAATACC 480
Db 9917 ACCCTCAAGCAACCAAAAAAGATCCAAACCTCAAAACCACTAAATCAAGGAATACC 9858
OY 481 ACACCAAGCCACAGAGAGAGCCACACATCAACACCAACCAAAACCAATCACTAC 540
Db 9857 ACACCAAGCCACAGAGAGAGCCACACATCAACACCAACCAAAACCAATCACTAC 9798
OY 541 CTGCTACCAACACACACAGAGAAATCCAAATCTCAAGTCAATGGAACCTTCAC 600
Db 9797 CTACTACCTCCACACACAGAGAAATCCAGAACTCAAGTCAATGGAACCTTCAC 9738
OY 601 TCACCTCTCCGAGAGCAATCTAGCCCTCTCAAGTCTCCACAACTCCAGACCA 660
Db 9737 TCACCTCTCCGAGAGCAATCCAGCCCTTCTCAAGTCTCTCAACATCCGAGTACCA 9678
OY 661 TCACACCTCTATCTCCACCCACACACAGCCAGTAGTATTAAAAAAA 714
Db 9677 TCACACCTCTATCTCCACCCACACACAGCCAGTAGTATTAAAAACATA 9624
RESULT 14
US-08-892-403A-1
; Sequence 1, Application US/08892403A
; Patent No. 5993824
; GENERAL INFORMATION:
; APPLICANT: Murphy, Brian R.
; APPLICANT: Collins, Peter L.
; APPLICANT: Whitehead, Stephen S.
; APPLICANT: Bukreyev, Alexander A.
; APPLICANT: Juhasz, Katalin
; TITLE OF INVENTION: PRODUCTION OF ATTENUATED RESPIRATORY
; TITLE OF INVENTION: SYNCYTIAL VIRUS VACCINES FROM CLONED NUCLEOTIDE SEQUENCES
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER REPROducible FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/892,403A
; FILING DATE: 15-JUL-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/047,634
; FILING DATE: 23-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/046,141
; FILING DATE: 09-MAY-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/021,773
; FILING DATE: 15-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 17634-000510
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15223 base pairs
; TYPE: nucleic acid

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 US-08-892-403A-1

Query Match 91.6%; Score 654.8; DB 2; Length 15223;
 Best Local Similarity 94.8%; Pred. No. 8.5e-169;
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAGTCACTACACTGCAATCATATGCAAGATGCAAGCAAGCCAGATCAAGACACA 60
 |||||||
 Db 4887 CACAAGTCACTACACTGCAATCATATGCAAGATGCAAGCAAGCCAGATCAAGACACA 4946
 QY 61 ACCCCAATACACTGCAATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 120
 |||||||
 Db 4947 ACCCCAATACACTGCAATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 5006
 QY 121 ATTACATGCAAGCAAGCCAGATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 180
 |||||||
 Db 5007 ATTACATGCAAGCAAGCCAGATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 5066
 QY 181 CAACCCACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
 |||||||
 Db 5067 CAATCCACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5126
 QY 241 ACTACAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCC 300
 |||||||
 Db 5127 ACCACAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCC 5186
 QY 301 GTGTTAACTTTGATGCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 360
 |||||||
 Db 5187 GTGTTAACTTTGATGCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5246
 QY 361 AAAAGATACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420
 |||||||
 Db 5247 AAAAGATACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5306
 QY 421 ACCTTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 480
 |||||||
 Db 5307 ACCTTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5366
 QY 481 ACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540
 |||||||
 Db 5367 ACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5426
 QY 541 CTGCTCACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
 |||||||
 Db 5427 CTGCTCACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5486
 QY 601 TCAAGCTCTCGAAGGCAATCTAAGCCTTCTCAAGTCTCACAAGATCGAGCACCA 660
 |||||||
 Db 5487 TCAAGCTCTCGAAGGCAATCTAAGCCTTCTCAAGTCTCACAAGATCGAGCACCA 5546
 QY 661 TCACACCTCTCTCACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 714
 |||||||
 Db 5547 TCACACCTCTCTCACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5600

RESULT 15

US-08-720-132-1
 ; Sequence 1, Application US/08720132
 ; Patent No. 6264957
 ; GENERAL INFORMATION:
 ; APPLICANT: Collins, Peter L.
 ; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, 8th Floor
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: U.S.A.

ZIP: 94111-3834
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/720,132
 FILING DATE: 27-SEP-1996
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/007,083
 FILING DATE: 27-SEP-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Parmelee, Steven W.
 REGISTRATION NUMBER: 31,990
 TELECOMMUNICATION INFORMATION:
 REFERENCE/DOCKET NUMBER: 15280-250-1
 TELEPHONE: 206-467-9600
 TELEFAX: 415-576-0300
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15223 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: other
 US-08-720-132-1

Query Match 91.6%; Score 654.8; DB 4; Length 15223;
 Best Local Similarity 94.8%; Pred. No. 8.5e-169;
 Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAGTCACTACACTGCAATCATATGCAAGATGCAAGCAAGCCAGATCAAGACACA 60
 |||||||
 Db 4887 CACAAGTCACTACACTGCAATCATATGCAAGATGCAAGCAAGCCAGATCAAGACACA 4946
 QY 61 ACCCCAATACACTGCAATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 120
 |||||||
 Db 4947 ACCCCAATACACTGCAATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 5006
 QY 121 ATTACATGCAAGCAAGCCAGATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 180
 |||||||
 Db 5007 ATTACATGCAAGCAAGCCAGATGATGCTGCAAGTGAATGAGTTCTTCCATGCTGAA 5066
 QY 181 CAACCCACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
 |||||||
 Db 5067 CAATCCACAAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5126
 QY 241 ACTACAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCC 300
 |||||||
 Db 5127 ACTACAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCCAAAGCAAGCC 5186
 QY 301 GTGTTAACTTTGATGCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 360
 |||||||
 Db 5187 GTGTTAACTTTGATGCTGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5246
 QY 361 AAAAGATACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 420
 |||||||
 Db 5247 AAAAGATACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5306
 QY 421 ACCTTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 480
 |||||||
 Db 5307 ACCTTAAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5366
 QY 481 ACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 540
 |||||||
 Db 5367 ACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5426
 QY 541 CTGCTCACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 600
 |||||||
 Db 5427 CTGCTCACCACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5486

Oy 601 TCAACCTCCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCCA 660
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 5487 TCAACCTCCTCGAAGCAATCTAAGCCCTTCTCAAGTCTCCACACATCCGAGACCCCA 5546
Oy 661 TCACAACCCCTCATCTTCACACCAACACACAGCCAGTAGTTATTAAAAAAA 714
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 5547 TCACAACCTCATCTTCACACCAACACACAGCCAGTAGTTATTAAAAACATA 5600

Search completed: May 1, 2003, 12:35:31
Job time : 53.1743 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 11:59:58 ; Search time 73.4679 Seconds
(Without alignments)
11499.705 Million cell updates/sec

Title: US-09-462-816-3
Perfect score: 715
Sequence: 1 cacaagaatcacactaacac.....gtagtattataaaaaaaa 715

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 746064 seqs, 590810554 residues

Total number of hits satisfying chosen parameters: 1492128

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCF_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	654.8	91.6	15223	9 US-09-847-173-1	Sequence 1, Appli
2	445.8	62.3	696	10 US-09-844-645-4	Sequence 4, Appli
3	83.6	11.7	793	10 US-09-878-574-4304	Sequence 4304, Ap
4	76.6	10.7	576	10 US-09-878-574-4286	Sequence 4296, Ap
5	76.4	10.7	529	10 US-09-883-965-2109	Sequence 2109, Ap
6	75.6	10.6	1030	10 US-09-878-574-4306	Sequence 4306, Ap
7	75	10.5	577	10 US-09-878-574-4312	Sequence 4312, Ap
8	74	10.3	639	10 US-09-878-574-4316	Sequence 4316, Ap
9	71.4	10.3	1075	10 US-09-864-761-19241	Sequence 19241, A
10	71.4	10.0	1403	10 US-09-864-761-2513	Sequence 2513, Ap
11	71.2	10.0	1635	10 US-09-864-761-20241	Sequence 20241, A
12	71.2	10.0	1973	10 US-09-864-761-3471	Sequence 3471, Ap
13	70	9.8	576	10 US-09-864-761-26582	Sequence 26582, A
14	68.6	9.6	1423	9 US-10-001-857-100	Sequence 100, App
15	68.2	9.5	491	10 US-09-878-574-4301	Sequence 4301, Ap
16	67.6	9.5	584	10 US-09-878-574-4310	Sequence 4310, Ap
17	67.2	9.4	1428	9 US-10-001-857-79	Sequence 79, Appl
18	63.6	8.9	555	10 US-09-864-761-9605	Sequence 9605, Ap
19	63.4	8.9	766	10 US-09-878-574-4344	Sequence 4344, Ap

C 20	63	8.8	528	10	US-09-878-574-4305	Sequence 4305, Ap
C 21	63	8.8	1267	12	US-10-001-843-45	Sequence 45, Appl
C 22	61.2	8.6	446	10	US-09-960-352-3400	Sequence 3400, Ap
C 23	61	8.5	448	10	US-09-878-574-4308	Sequence 4308, Ap
C 24	60.2	8.4	341	10	US-09-960-352-12302	Sequence 12302, A
C 25	59	8.3	545	10	US-09-878-574-4299	Sequence 4299, Ap
C 26	57.6	8.1	1557	9	US-10-001-873-1	Sequence 1, Appl1
C 27	57.4	8.0	425	10	US-09-960-352-4010	Sequence 4010, Ap
C 28	57.4	8.0	1938	9	US-09-834-975-1039	Sequence 1039, Ap
C 29	57.2	8.0	11956	9	US-10-270-336-3	Sequence 3, Appl1
C 30	57	8.0	446	10	US-09-864-761-20699	Sequence 20699, A
C 31	56	7.8	155074	9	US-10-026-188-6	Sequence 6, Appl1
C 32	55.6	7.8	283	10	US-09-960-352-9095	Sequence 9095, Ap
C 33	54	7.6	415	10	US-09-960-352-2223	Sequence 2223, Ap
C 34	53	7.4	15720	9	US-10-025-380-1058	Sequence 1058, Ap
C 35	53	7.4	15720	10	US-09-922-217-1058	Sequence 1058, Ap
C 36	53	7.4	15720	10	US-09-833-263-1058	Sequence 1058, Ap
C 37	51.8	7.2	479	10	US-09-864-761-768	Sequence 768, Ap
C 38	51.8	7.2	830	10	US-09-864-761-19531	Sequence 19531, A
C 39	51.4	7.2	396	9	US-09-970-966-53	Sequence 53, Appl
C 40	51.4	7.2	396	9	US-09-825-294-53	Sequence 53, Appl
C 41	51.4	7.2	462	9	US-10-083-357-341	Sequence 341, Appl
C 42	51.2	7.2	463	10	US-09-960-352-7186	Sequence 7186, Ap
C 43	51	7.1	600	10	US-09-864-761-6690	Sequence 6690, Ap
C 44	50.8	7.1	439	10	US-09-960-352-1009	Sequence 1009, Ap
C 45	50.8	7.1	516	10	US-09-960-352-5785	Sequence 5785, Ap

ALIGNMENTS

RESULT 1
US-09-847-173-1
; Sequence 1, Application US/09847173
; Publication No. US20020182228A1
; GENERAL INFORMATION:
; APPLICANT: Collins, Peter L.
; TITLE OF INVENTION: PRODUCTION OF INFECTIOUS RESPIRATORY
; SYNCTIAL VIRUS FROM CLONED NUCLEOTIDE SEQUENCES
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 94111-3634
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.2;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/847,173
; FILING DATE: 03-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/720,132
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 15280-250-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-467-9600
; TELEFAX: 415-576-0300
; INFORMATION FOR SEQ ID NO. 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15223 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other

SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-847-173-1

Query Match 91.6%; Score 654.8; DB 9; Length 15223;
Best Local Similarity 94.8%; Pred. No. 1.4e-164;
Matches 677; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

QY 1 CACAAGTCACTACATGCAATCATGCAATGCAACAGCCAGATCAAGACACA 60
DB 4887 CACAAGTCACTACATGCAATCATGCAATGCAACAGCCAGATCAAGACACA 4946
QY 61 ACCCAACATACCTACTAGATGATCTGAGTGGATGAGTCTGCAATCTGTGAA 120
DB 4947 ACCCAACATACCTACTAGATGATCTGAGTGGATGAGTCTGCAATCTGTGAA 5006
QY 121 ATTACATCAACACACACACATGATGATGCAACACACAGATGCAATGCAACAC 180
DB 5007 ATTACATCAACACACACACATGATGATGCAACACACAGATGCAATGCAACAC 5066
QY 181 CAACCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 240
DB 5067 CAATCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 5126
QY 241 ACTACAAACACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 300
DB 5127 ACCCAACACAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5186
QY 301 GTGTTTACCTTGTACCTGACATATGAGCAACATGCAACCTGAGGCTATCTGC 360
DB 5187 GTGTTTACCTTGTACCTGACATATGAGCAACATGCAACCTGAGGCTATCTGC 5246
QY 361 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 420
DB 5247 AAAAGATATCCAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5306
QY 421 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 480
DB 5307 ACCTTCAAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5366
QY 481 ACCACCAAGCCACAGAGAGCAACCAACCAACCAACCAACCAACCAACCAAC 540
DB 5367 ACCACCAAGCCACAGAGAGCAACCAACCAACCAACCAACCAACCAACCAAC 5426
QY 541 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 600
DB 5427 CTGCTCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 5486
QY 601 TCAACCTCTCGAAGGCAATGAGCCCTTCAAGTCTCCACACATCCGAGCACA 660
DB 5487 TCAACCTCTCGAAGGCAATGAGCCCTTCAAGTCTCCACACATCCGAGCACA 5546
QY 661 TCAACCTCTCGAAGGCAATGAGCCCTTCAAGTCTCCACACATCCGAGCACA 714
DB 5547 TCAACCTCTCGAAGGCAATGAGCCCTTCAAGTCTCCACACATCCGAGCACA 5600

RESULT 2
US-09-844-645-4

Sequence 4, Application US/09844645
Patent No. US20020102242A1

GENERAL INFORMATION:

APPLICANT: Biles, David E.

McDaniel, Larry S.

Curtiel, David T.

TITLE OF INVENTION: COMPOSITION AND METHODS FOR
ADMINISTERING PNEUMOCOCCAL DNA

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: 530 Fifth Avenue

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/844,645

FILING DATE: 27-Apr-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/759,505

FILING DATE: 04-DEC-1996

ATTORNEY/AGENT INFORMATION:

NAME: Frommer, William S.

REGISTRATION NUMBER: 25,506

REFERENCE/DOCKET NUMBER: 454312-2450

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 840-3333

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 696 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-844-645-4

Query Match 62.3%; Score 445.8; DB 10; Length 696;
Best Local Similarity 95.3%; Pred. No. 2.2e-109;
Matches 470; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

QY 1 CACAAGTCACTACATGCAATCATGCAATGCAACAGCCAGATCAAGACACA 60
DB 199 CACAAGTCACTACATGCAATCATGCAATGCAACAGCCAGATCAAGACACA 258
QY 61 ACCCAACATACCTACTAGATGATCTGAGTGGATGAGTCTGCAATCTGTGAA 120
DB 259 ACCCAACATACCTACTAGATGATCTGAGTGGATGAGTCTGCAATCTGTGAA 318
QY 121 ATTACATCAACACACACATGATGCAACACACAGATGCAATGCAACACAC 180
DB 319 ATTACATCAACACACACATGATGCAACACACAGATGCAATGCAACACAC 378
QY 181 CAACCAACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 239
DB 379 CAATCCACACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 438
QY 240 CACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 299
DB 439 CACTACAAACCAACGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 498
QY 300 AGTGTTCACCTTGTACCTGAGATGAGCAACCAACCAACCAACCAACCAAC 359
DB 499 AGTGTTCACCTTGTACCTGAGATGAGCAACCAACCAACCAACCAACCAAC 558
QY 360 CAAAAGATATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 419
DB 559 CAAAAGATATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 618
QY 420 AACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 479
DB 619 AACCTTCAAGACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 678
QY 480 CACGACCAAGGCC 492
DB 679 CACGACCAAGGCC 691

RESULT 3

US-09-878-574-4304/C
Sequence 4304, Application US/09878574


```

; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4304
; LENGTH: 793
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(793)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-014-Q1-B1-C11
US-09-878-574-4304

```

```

Query Match          11.7%; Score 83.6; DB 10; Length 793;
Best Local Similarity 48.4%; Pred. No. 1.2e-12;
Matches 276; Conservative 0; Mismatches 292; Indels 2; Gaps 2;

```

```

QY 1 CACAAAGTCACACTAACCACTGCAATCTATACAGATGCCAACAGGAGATCAAGACACA 60
DB 700 CACCAAAANACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 641
QY 61 ACCCCA-ACATACCTCAGTCCAGATCCGATGTTGGATTCAGCTTCTCCATCTCTCTGA 119
DB 640 AACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 581
QY 120 AATTACATCACAAC-CACACGCAATAGCTTCAACAACCAAGAGTCAAGTCAAC 178
DB 580 AACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 521
QY 179 TGCAACCCCAACAGTCAAGCTTAAACACACACCAACCAACCAACCAACCAACCAACCA 238
DB 520 CACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 461
QY 239 CCACCTACCAAAACAGCCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCA 298
DB 460 ACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 401
QY 299 AAGTGTAACTTTGACCTTGACGATATGCGACACCAATTCAACTGCTGGGCTATCT 358
DB 400 ACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 341
QY 359 GCAAAAGATACCAAAACCAAAACCAAGAAAGAAACCAACCAACCAACCAACCAACCAAC 418
DB 340 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 281
QY 419 CAACCTTCAAGACCAACCAAAAGATCTCAAACTTCAAACTCAAACTCAAAAGCAAGTAC 478
DB 280 AAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 221
QY 479 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 538
DB 220 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 161
QY 539 CACTGCTCACCACACACCAACCAACCAAGCAAAATC 568
DB 160 CACCAACCAACCAACCAACCAACCAACCAACCAAC 131

```

```

RESULT 4
US-09-878-574-4296/C
; Sequence 4296, Application US/09878574
; Patent No. US20020110548A1

```

```

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4296
; LENGTH: 576
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(576)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: LIB3028-013-Q1-B1-G1
US-09-878-574-4296

```

```

Query Match          10.7%; Score 76.6; DB 10; Length 576;
Best Local Similarity 50.9%; Pred. No. 7.4e-11;
Matches 207; Conservative 0; Mismatches 199; Indels 1; Gaps 1;

```

```

QY 186 CACACAGTCACAGCTAAACACACACACCAACCAACCAACCAACCAACCAACCAACCA 244
DB 556 CACCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 497
QY 245 CAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 304
DB 496 ACAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 437
QY 305 TTAACCTTGTACCTGCGATATGCGAACAATCCGATCTGCTGCTATCTGCAAAA 364
DB 436 CCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 377
QY 365 GAATACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 424
DB 376 AACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 317
QY 425 TCAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 484
DB 316 ACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 257
QY 485 CCAAGCCCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCT 544
DB 256 CAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 197
QY 545 TCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCTCA 591
DB 196 CCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 150

```

```

RESULT 5
US-09-983-965-2109/C
; Sequence 2109, Application US/09983965
; Patent No. US20020137160A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678

```

```

RESULT 6
US-09-878-574-4306/C
; Sequence 4306, Application US/09878574
; Patent No. US20020110546A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 4306
; LENGTH: 1030
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1030)

```

```

? RESULT 7
? US-09-878-574-4312/c
? Sequence 4312, Application US/09878574
? Patent No. US20020110548A1
?
? GENERAL INFORMATION:
?
? APPLICANT: Byrum, Joseph R.
? APPLICANT: La Rosa, Thomas J.
? APPLICANT: Thompson, Michael D.
? TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
? TITLE OF INVENTION: Plants
? FILE REFERENCE: 38-21(15401)B
?
? CURRENT APPLICATION NUMBER: US/09/878,574
?
? CURRENT FILING DATE: 2001-12-21
?
? PRIOR APPLICATION NUMBER: 09/333,535
?
? PRIOR FILING DATE: 1999-06-14
?
? NUMBER OF SEQ ID NOS: 15775
?
? SEQ ID NO 4312
?
? LENGTH: 577
?
? TYPE: DNA
?
? ORGANISM: Glycine max
?
? FEATURE:
?
? NAME/KEY: unsure
?
? LOCATION: (1)..(577)
?
? OTHER INFORMATION: unsure at all n locations
?

```



```
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO: 19241
;; LENGTH: 1075
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL078472.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
;; OTHER INFORMATION: EST_HUMAN HIT: AV739739.1, EVALUATE 1.00e+00
;; OTHER INFORMATION: NT HIT: AL163201.2, EVALUATE 2.00e-19
US-09-864-761-19241
```

```
Query Match          10.0%  Score 71.4:  DB 10:  Length 1075:
Best Local Similarity 45.7%:  Pred. No. 2.4e-09:
Matches 249:  Conservative 0:  Mismatches 296:  Indels 0:  Gaps 0:
```

```
QY 16 ACAACTGCAATCATACAGATGCAACAAGCCAGATCAAGAACACACCCCAATACCTC 75
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 711 ACCACCAACCATCACCACACCAACCATCACCACCAACCATCACCACCAACCATCACC 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 76 ACTCAGAGATCCAGCTTGATGAGTCTTCCAGTCTTCCAGTCTTCCAGTCTTCCAGTCT 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 651 ACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 ACCACCATATAGTCTTCAACAACCAACCAACCAACCAACCAACCAACCAACCAAC 195
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 591 ACCACCATATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 AAGACTAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAGC 255
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 531 ACCACCAACCATCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 CAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 471 ACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 412
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 CCCTGAGCATATGCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 375
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 411 ACCAGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 AAAAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 435
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 351 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 292
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 436 AAAAAAGATCTCAAAACCTCAAAACCTCAAAACCTCAAAACCTCAAAACCTCAAAAC 495
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 291 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 496 GAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 555
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 231 ACCATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 556 ACCAC 560
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 171 ACCAC 167
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 10
US-09-864-761-2513/c
```

```
;; Sequence 2513, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharron G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
;; FILE REFERENCE: Aemica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
```

```
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO: 2513
;; LENGTH: 1403
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL078472.1
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 27
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 19
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 34
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.1
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 28
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 43
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 20
;; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 25
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 18
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 21
US-09-864-761-2513
```

```
Query Match          10.0%  Score 71.4:  DB 10:  Length 1403:
Best Local Similarity 45.7%:  Pred. No. 2.7e-09:
Matches 249:  Conservative 0:  Mismatches 296:  Indels 0:  Gaps 0:
```

```
QY 16 ACAACTGCAATCATACAGATGCAACAAGCCAGATCAAGAACACCAACCAATACCTC 75
```

```

Db 1061 ACCACCACCATCACCACCACCAACCATCAGCACACCATCACCACCAACCATCACC 1002
QY 76 ACTCAGATTCCTCAGCTTGATCGCTTCCTCCATCTGCTGAAATTACATCAAAAC 135
Db 1001 ACTACACCAACACCAACCAACCATCAGCACACCATCACCACCAACCATCACCAC 942
QY 136 ACCACCATCTAGTCTTCAACAAACACAGAGTCAAGTCAAACTGCAACCCCAACAGTC 195
Db 941 ACCACCATCTAGCCTACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 882
QY 136 AAGACTAAAAACACAAACAACCAACCAACCAACCAACCAACCAACCAACCAAC 255
Db 881 ACCACCATCTAGCCTACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 822
QY 256 CAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 315
Db 821 ACTACCATCTAGCCTACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 762
QY 316 CCCTGACGATATGCGAGCAACATCCAGCTGCTGGCTATCTGCAAAAGATACCAAC 375
Db 761 ACCACCATCTAGCCTACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 702
QY 376 AAAAAACCAAGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 435
Db 701 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 642
QY 436 AAAAAAGATCTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAACTCAAA 495
Db 641 ATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 582
QY 496 GAAGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 555
Db 581 ACCATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 522
QY 556 ACCAC 560
Db 521 ACCAC 517

```

RESULT 11

```

US-09-864-761-20241
: Sequence 20241 Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: FILE REFERENCE: Acomlca-X-1
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665

```

```

: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 20241
: LENGTH: 1635
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO AC006547.9
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
: OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 8,6
: OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 9
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 12
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8,4
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-20241

```

```

Query Match 10.0%; Score 71.2; DB 10; Length 1635;
Best Local Similarity 46.6%; Pred. No. 3.3e-09;
Matches 261; Conservative 0; Mismatches 298; Indels 1; Gaps 1;

```

```

QY 133 ACCACCACTTACTAGCTTAAACAACCAAGGTAAGTCAAACTGCAACCCACAACA 192
Db 794 ACCACCACTTACTAGCTTAAACAACCAAGGTAAGTCAAACTGCAACCCACAACA 853
QY 193 GTCAGACTAAACCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 252
Db 854 ACCACCACTTACTAGCTTAAACAACCAAGGTAAGTCAAACTGCAACCCACAACA 913
QY 253 CGCCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 312
Db 914 ACCACCACTTACTAGCTTAAACAACCAAGGTAAGTCAAACTGCAACCCACAACA 973
QY 313 GTACCTGCGAGCATATGACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 372
Db 974 ACCATCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1033
QY 373 AAAAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 432
Db 1034 ATCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1093
QY 433 AAAAAAACCAAGATCTCAAACTTAAACCAACCAACCAACCAACCAACCAACCAAC 492
Db 1094 ACCACCACTTACTAGCTTAAACAACCAAGGTAAGTCAAACTGCAACCCACAACA 1153
QY 493 ACAGAAAGGCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 552
Db 1154 ATCACTACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1213
QY 553 AAAAAAACCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 612
Db 1214 ACTACCACTTACTAGCTTAAACAACCAAGGTAAGTCAAACTGCAACCCACAACA 1273

```

OY	613	GAAGGCAATTAAGCCCTTCTCAAGTCTGCACA	-CATCCGAGGCAACCATCAACCTTC	671
Db	1274	ACAACCAACATCACTACCCACATCACCACATACCA	CGCATCTACTACCAACCAACCAACCAAC	1333
OY	672	ATCTCCACCCCAACACAAACAC		691
Db	1334	ATCACCAATCACTTCCACCAAC		1353

RESULT 12

```

Sequence 347 Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
  APPLICANT: Penn, Sharon G.
  APPLICANT: Rank, David R.
  APPLICANT: Hanzel, David K.
  APPLICANT: Chen, Wensheng
  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
  TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
  FILE REFERENCE: Aeomica-X-1
  CURRENT APPLICATION NUMBER: US/09/864,761
  CURRENT FILING DATE: 2001-05-23
  PRIOR APPLICATION NUMBER: US 60/180,312
  PRIOR FILING DATE: 2000-02-04
  PRIOR APPLICATION NUMBER: US 60/207,456
  PRIOR FILING DATE: 2000-05-26
  PRIOR APPLICATION NUMBER: US 09/632,366
  PRIOR FILING DATE: 2000-08-03
  PRIOR APPLICATION NUMBER: GB 24263,6
  PRIOR FILING DATE: 2000-10-04
  PRIOR APPLICATION NUMBER: US 60/236,359
  PRIOR FILING DATE: 2000-09-27
  PRIOR APPLICATION NUMBER: PCT/US01/00666
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00667
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00664
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00665
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00665
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00668
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00663
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00662
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00661
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: PCT/US01/00670
  PRIOR FILING DATE: 2001-01-30
  PRIOR APPLICATION NUMBER: US 60/234,687
  PRIOR FILING DATE: 2000-09-21
  PRIOR APPLICATION NUMBER: US 09/608,408
  PRIOR FILING DATE: 2000-06-30
  PRIOR APPLICATION NUMBER: US 09/774,203
  PRIOR FILING DATE: 2001-01-29
  NUMBER OF SEQ ID NOS: 49117
  SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
  SEQ ID NO 3471
  LENGTH: 1973
  TYPE: DNA
  ORGANISM: Homo sapiens
  FEATURE:
    OTHER INFORMATION: MAP TO AC006547.9
    OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 17
    OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 11
    OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8,6
    OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9
    OTHER INFORMATION: EXPRESSED IN HL10C, SIGNAL = 12
    OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 11

```

```

; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 12
US-09-864-761-3471

```

Query Match	10.0%;	Score 71.2;	DB 10;	Length 1973;
Best Local Similarity	46.6%;	Pred. No. 3.6e-09;		
Matches 261;	Conservative	0;	Mismatches 298;	Indels 1;
				Gaps 1.

[illegible]

RESULT 13

```

US-09-864-761-26582/C
: Sequence 26582, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
: FILE REFERENCE: Aecmca-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04

```

```

? APPLICANT: Chen, Sei-Yu
? APPLICANT: Sun, Yongming
? APPLICANT: Liu, Chenghua
? TITLE OF INVENTION: Compositions and Methods Relating to Lung Specific Genes and F
? FILE REFERENCE: DEX-0273
? CURRENT APPLICATION NUMBER: US/10/001,857
? CURRENT FILING DATE: 2001-11-20
? PRIOR APPLICATION NUMBER: 60/252,054
? PRIOR FILING DATE: 2000-11-20
? NUMBER OF SEQ ID NOS: 208
? SOFTWARE: PatentIn version 3.1
? SEQ ID NO 100
? LENGTH: 1423
? TYPE: DNA
? ORGANISM: Homo sapien
? US-10-001-857-100

```

QY	118	GAATTTACATCAAAACACACCACCCTATAGTTCAACCAACACAGAGAGTCAATGCTAAC	177
Db	930	GAATTTCAACACACCACTAAATTCACCACTAACTAAACAAAAACAAAAAACCAACAA	989
QY	178	CTGCACCCCAACAGTCAAGACTAAACACACAAACACACCCCAACACACCCAGCAAG	237
Db	990	CACAAAAAGAAAAAATTAATAAACCAACATCAACCAACAAACCAAAAAACAAACA	1049
QY	238	CCCACTCAAAACACAGCCCAAAACCAACCAACCAAAACCAATATATGATTTCACTTC	297
Db	1050	AAAAATACAAAAAATCATTAATAAAAAACAAAAAACACATACATATATATATATATAT	1109
QY	298	GAAGCTTTAACTTTGTACCCCTGCACATATGAGCAACATTCCAACCTGCTGGCTATC	357
Db	1110	CAAAACAACCAACAAACAAACACACACACATATAGCATATACCAAAATATAAAAACAA	1169
QY	358	TGCAAAAGATATCCCAAAACAAAAACCAAGAAAAAACCAACCAACAGCTATCAAAAAA	417
Db	1170	ATTATATACAGAAACAACCTTACAAAAAATAATATATAAAAACATTAATAATATATATA	1229
QY	418	CCAACTTCAGAGCAACCAAAAAAGATCTCAAAACCTCAAAACACATTAACCAAGCACTA	477
Db	1230	AAACTATAAATATACAAAAATATTAATAATCTTAACCAACCAAAAAAATATACACAAAAA	1289
QY	478	CCCAACCAACGCCCACAGAGAGGCCAAGCCATCAACACCACCAAAACAAATATCACAACT	537
Db	1290	AACAAACCCAAACATTAATAAAAAACAAAAACAAACACAAAAAATAAAAAATATAAA	1349
QY	538	ACACTGCTCACCAACACACACAGAGAAATCCAAACCTCACAGTCAATAGCAATCTTC	597
Db	1350	AAAAACGTAAACGATTAATAAATAATTAATAAATAAATAAATAAATAAATAAATAAATA	1409
QY	598	CAC 600	
Db	1410	AAC 1412	

```

RESULT 15
US-09-878-574-4301/C
: Sequence 4301, Application US/09878574
: Patent No. US20020110548A1
: GENERAL INFORMATION:
: APPLICANT: Byrum, Joseph R.
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Thompson, Michael D.
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(15401)B
: CURRENT APPLICATION NUMBER: US/09/878,574
: CURRENT FILING DATE: 2001-12-21
: PRIOR APPLICATION NUMBER: 09/333,535
: PRIOR FILING DATE: 1999-06-14

```


GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 1, 2003, 10:08:32 ; Search time 884.239 Seconds
(Without alignments)
13095.757 Million cell updates/sec

Title: US-09-462-816-3
Perfect score: 715
Sequence: 1 cacaaagtcactaacac.....glagtattcaaaaaaaaaa 715

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 16154066 segs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estba: *
2: em_esthum: *
3: em_estlin: *
4: em_estmu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hlc: *
9: gb_est1: *
10: gb_est2: *
11: gb_hlc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: gb_gss: *
18: em_gss_hum: *
19: em_gss_inv: *
20: em_gss_pla: *
21: em_gss_vit: *
22: em_gss_fun: *
23: em_gss_man: *
24: em_gss_mus: *
25: em_gss_other: *
26: em_gss_pro: *
27: em_gss_rtd: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	120	16.8	895	17	CNS0071A
2	104.6	14.6	946	17	AG127412
3	103.2	14.4	922	17	CNS0073W
4	101.8	14.2	919	17	CNS04EMX
5	100	14.0	1201	9	AL581589
6	99	13.8	1101	17	CNS00EXE
					AL071370 Drosophil

C	7	95.8	13.4	1097	17	CNS01224	AL102202 Drosophil
	8	95	13.3	791	17	CNS009KS	AL053801 Drosophil
	9	95	13.3	1101	17	CNS00LOO	AL068607 Drosophil
	10	94.4	13.2	1216	17	AG135357	AG135357 Pan trogl
	11	94.2	13.2	815	17	B12686	B12686 F27J14-T7 I
	12	93.8	13.1	1008	17	AG137085	-S137085 Pan trogl
	13	93.6	13.1	884	17	CNS006U0	AL063923 Drosophil
	14	93.6	13.1	969	17	AG0743309	AG0743309 HS_5387_B
	15	93.2	13.0	1141	17	AG0743305	AG0743305 HS_5387_B
	16	92.8	13.0	981	17	AG127518	AG127518 Pan trogl
	17	92.6	13.0	1183	17	AG136828	AG136828 Pan trogl
	18	92.4	12.9	859	17	AG128925	AG128925 Pan trogl
	19	92.4	12.9	880	17	AG139490	AG139490 Pan trogl
	20	92	12.9	1137	17	AG0743326	AG0743326 HS_5387_B
	21	92	12.9	1225	17	CNS0166K	AL106358 Drosophil
	22	91.4	12.8	700	17	AG127423	AG127423 Pan trogl
	23	91.4	12.8	885	17	CNS031TM	AL249907 Tetrarodon
	24	91	12.7	1099	17	AG0743360	AG0743360 HS_5387_B
	25	90.8	12.7	1081	17	AG135328	AG135328 Pan trogl
	26	90.6	12.7	1491	12	BE882936	BE882936 601505836
	27	90.2	12.6	866	17	AG126308	AG126308 Pan trogl
	28	90	12.6	953	17	CNS006PF	AL063557 Drosophil
	29	90	12.6	1101	17	CNS017FC	AL107970 Drosophil
	30	89.8	12.6	896	17	AG141027	AG141027 Pan trogl
	31	89.4	12.5	909	17	CNS00JTL	AL076720 Drosophil
	32	89.4	12.5	955	17	AG076494	AG076494 Pan trogl
	33	89.4	12.5	1024	17	AG133080	AG133080 Pan trogl
	34	89.4	12.5	1054	17	AG137135	AG137135 Pan trogl
	35	89.4	12.5	1188	17	AG135332	AG135332 Pan trogl
	36	89.4	12.5	1353	17	AG128010	AG128010 Pan trogl
	37	89	12.4	1101	17	CNS00JTL2	AL076720 Drosophil
	38	89	12.4	1144	17	AG0743364	AG0743364 HS_5387_B
	39	88.8	12.4	938	17	AG132199	AG132199 Pan trogl
	40	88.4	12.4	926	17	AG078232	AG078232 HS_3176_B
	41	88.2	12.3	966	17	AG126022	AG126022 Pan trogl
	42	88.2	12.3	987	17	CNS00K3Q	AL077186 Drosophil
	43	88.2	12.3	1101	17	CNS010B3	AL098745 Drosophil
	44	88	12.3	860	17	CNS018FL	AL109275 Drosophil
	45	88	12.3	960	17	CNS005F3	AL059925 Drosophil

ALIGNMENTS

RESULT 1
CNS0071A/c 895 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR14809 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL066286
VERSION AL066286.1 GI:4945153
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 895)

AUTHORS Genoscope.
TITLE Direct Submissio
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr

COMMENT

- Web : www.genoscope.cns.fr
determination of this BAC end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Otsuoka and
Aaron Mammosser in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial


```
OY 615 AGGCATCTAAGCCCTTCTCAAGTCTCCACACATCCGACACCATCATCAACCCCTCATC 674
      ||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 753 CAGACACCCCAAAAAAACCCTCCACACAAAAAACAACACACACAAACACACCA 812.
OY 675 TCCACCCACACACACGCGAGTAGTATTAAAAA 714
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 813 CCAAAACCCACAGAAAACACACACCCCAACAA 852

RESULT 3
CNS0073W 922 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR14D09 of Rpci-98 library from Drosophila melanogaster (fruit
fly); genomic survey sequence.
ACCESSION AL066784.1 GI:4945247
VERSION AL066784
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Phytophagae; Drosophilidae; Drosophila.
1 (bases 1 to 922)
Genoscope.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL Bp 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACS. For further information
please see http://www.fruitfly.org/The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pletier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named Rpci-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source location/Qualifiers
1..922
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone_lib="BACR14D09"
/note="end : TET3"

BASE COUNT 223 a 95 c 109 g 221 t 274 others
ORIGIN
Query Match 14.4%; Score 103.2; DB 17; Length 922;
Best Local Similarity 24.8%; Pred. No. 5.2e-10;
Matches 102; Conservative 164; Mismatches 144; Indels 1; Gaps 1;
```

```
Db 742 AAAAAAACAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAAAMAAAAA 683
OY 421 ACCTTCACACACCAAAAAAGATGTCACCTCAACCTCAACCTCAACCTCAACCTCAACCT 480
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 682 MNCAMMMMMCAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMAAMMMMA 623
OY 481 ACCACCAACCCACAGAGAAGCCCAACATCAACACCAACCAACCAACCAACCAACCAACCA 540
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 622 AAMACAMAMACAAAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 563
OY 541 CTCTCACCACACACACACAGGAAATCCAAATCTCAAGTCAATGCA 591
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 562 AAMACAMAMACASAAARAAAMCMACCCMAAAAAAARAAAAA 512

RESULT 4
CNS04ENY 919 bp DNA linear GSS 21-MAY-2000
LOCUS Tetraodon nigroviridis genome survey sequence T7 end of clone
DEFINITION 104P14 of library G from Tetraodon nigroviridis, genomic survey
sequence.
ACCESSION AL287287
VERSION AL287287.1 GI:8025774
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphi; Acanthopterygii; Percomorpha; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 919)
REFERENCE
AUTHORS Roest-Elizola, H., Jallion, O., Dasilva, C., Bouneau, I., Fisher, C.,
Saurin, W., and Weissenbach, J.
Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
Unpublished
2 (bases 1 to 919)
REFERENCE
AUTHORS Roest-Elizola, H., Jallion, O., Dasilva, C., Fisher, C.,
Bouneau, I., Billault, A., Queller, F., Saurin, W., Bernot, A. and
Weissenbach, J.
Characterization and repeat analysis of the compact genome of the
freshwater puterfish Tetraodon nigroviridis
Unpublished
3 (bases 1 to 919)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.

FEATURES
source location/Qualifiers
1..919
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone_lib="104P14"
/note="Genoscope sequence ID : C08G104Dh07L1-end : T7"

BASE COUNT 99 a 86 c 254 g 409 t 71 others
ORIGIN
Query Match 14.2%; Score 101.8; DB 17; Length: 919;
Best Local Similarity 46.0%; Pred. No. 9.6e-10;
Matches 227; Conservative 36; Mismatches 222; Indels 9; Gaps 1;
```

QY 239 GCACGACCAACCAAGCCCAAAACCAACCAACCAAAACCAATGATTTTCAGCTTG 298
 Db 777 CAACCCMACCCACCAACCAACACACAMCCCCCAACACCAACAA-----AAACAM 727
 QY 299 AAGTGTTTAACTTTGTACCTTGCAAGCATATGACGACAGCAATGCCAAGCTGTGGGCTATCT 358
 Db 726 AAACACAAACCCCAAAAAAAAAAAAAAAAAAACACAAACACCAAAAAAAAAAAAAACAAACAAAA 667
 QY 359 GCACAAAGATACCAACCAAAAAAAAAACGAGAAAGAAACCAACCAAGCTTACAAAAAAC 418
 Db 666 AAAAAACAAACACACCAACCAACACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 607
 QY 419 CAACCTTAAAGCAACCAACCAAAAGTCTCAAAACCTCAACCAACCACTAAACCAAGAAAGTAC 478
 Db 606 ACACAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 547
 QY 479 CCACCCACCAAGCCACAGAGAGCCCAACCATCAACACACCAACCAACCAACCAACCAACCACTA 538
 Db 546 NCMCCACACACACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 487
 QY 539 CACTGCTCAACCAACACACACACAGGAAATCCAAAACTCAAGTCAAAATGGAAACCTTCC 598
 Db 486 AAAAAAAMAMAMACACACAAACCAACACACACACACACACACACACACACACACACACACAC 427
 QY 599 ACTCAACCTCCTCC 612
 Db 426 CCNCCACACCCCC 413

RESULT 5						
AL581589/c						
LOCUS	AL581589	1201 bp	mrna	linear	EST 16-FEB-2001	
DEFINITION	AL581589 LTI_FLO11_BCI	Homo sapiens	CDNA clone	CS0DG004YF15	3 prime	
	/ mRNA sequence.					

ACCESSION	AL581589	:
VERSION	AL581589.1	GI:12948741
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	

REFERENCE	1 (bases 1 to 1201)
AUTHORS	Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope

FEATURES
source
Location/Qualifiers
1. .1201

```

/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

```

BASE COUNT	92 a	57 c	490 g	461 t	101 others
ORIGIN					
Query Match		14.0%;	Score 100;	DB 9;	Length 1201;
Best Local Similarity		45.5%;	Pred. No. 2e-09;		

	Matches	256; Conservative	39;	Mismatches	264;	Indels	4;	Gaps	2;
OY	132	AACCACCACCATCTACTGCTTCAACACACACGAGCTCAAGTCAAGTCAAACTGCAACCCACAC	191						
Db	1152	AAAAAAAAACCAAA	1093						
OY	192	AGTCAAGACTAAATAAACACAAACACCAACCAACCAACCAACCAACCAACCAACCAACCA	251						
Db	1092	AACCCNCAAC	1033						
OY	252	ACGCCAAACCAACACCAACCAACCAACCAATTAATGATTTTCACTTGCAGAGTGTTTACTT	311						
Db	1032	ACCCCAACCAAA	973						
OY	312	TGTACCTGTGAGATATGAGCAACAATCCAGCTGCTGGGCTATCTGCAAAAGAAATATAC	371						
Db	972	AAACACAAACCCCAAAAAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	913						
OY	372	AAACCAAAAAAAAAACGAAAGAAAGAAACCAACCAACCAACCAACCAACCAACCAACCAAC	431						
Db	912	MMCCMCAACCAACCAAAAAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	855						
OY	432	AACCAAAAAAAAAAGATCTCAAACTCAAAACCACTTAACCAAGAGAGTACCACACCAACGC	491						
Db	854	AAC - AACCAACACCCCAACCAAAAAAAAAACCAAAAAAAAAACCAACCAACCAACCAAC	797						
OY	492	CACGAAAGGACCAACCATCAACACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	551						
Db	796	AAAAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	737						
OY	552	CAACACCACAGAAATCCAAAACCTCACAAGTCAATGAAATCTTCACATCAACCTCTCTC	611						
Db	736	MAAACACCCCAACCAAAAAAAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	677						
OY	612	CGAAGCAATCTAAGCCCTTCTCAAGTCTCACAACATCGAGACCCATCAACACCTTC	671						
Db	676	CACCCCCCTCACACCCCTCCCTCCACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	617						
OY	672	ATGCCACCAACACACACACGCC	694						
Db	616	CCCAACCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT	594						

RESULT	6
CNS00EFXE/c	
LOCUS	
DEFINITION	CNS00EFXE 1101 bp DNA linear GSS 03-JUN-1999
ACCESSION	Drosophila melanogaster genome survey sequence TET3 end of BAC:
VERSION	BACR32c1 of RPCL-98 library from Drosophila melanogaster (fruit
KEYWORDS	fly), genomic survey sequence.
SOURCE	AL071370
ORGANISM	AL071370.1 GI:4951210
	GSS.
	Drosophila melanogaster.
	Drosophila melanogaster.
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
	Ephydroidea; Drosophilidae; Drosophila.
	1 (bases 1 to 1101)
REFERENCE	Genoscope.
AUTHORS	Direct Submission
TITLE	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL	BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
COMMENT	- Web : www.genoscope.cns.fr) determination of this BAC-end sequence was carried out as part of a

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Séquençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org/TheBDGP/Drosophila>
melanogaster BAC library was prepared by Kazutoyo Oseegawa and
Aaron Mammoser in Plietier de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPc1-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the

isogenic strain y2; cn bw sp, the same strain used for the BDGP's pl and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
location/Qualifiers
1. 1101
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACR32C19"
/clone_lib="RPCL-98"
/note="end : TET3"

BASE COUNT 326 a 104 c 183 g 186 t 302 others
ORIGIN

Query Match 13.8%; Score 99; DB 17; Length 1101;
Best Local Similarity 25.6%; Pred. No. 3.2e-09;
Matches 104; Conservative 145; Mismatches 157; Indels 0; Gaps 0;

```
OY 153 AACACACGAGAGTCAAGTCAACCTGCAACCCACACACTCAAGATAAACAACACAC 212
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1081 AAAMCMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 1022
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 213 AACACACGAGAGTCAAGTCAACCTGCAACCCACACACTCAAGATAAACAACACAC 272
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1021 AAHAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 962
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 273 CAACCCCAATATGATTTTCATCTGAGAGTGTATTAATCTTACCCTGCACATATGAG 332
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 961 MAHAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 902
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 333 CAACATCAACCTGCTGCTGCTATCTGCAAAAGATACCAACCAAAACAGGAAAGAA 392
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 901 CMAAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 842
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 393 AACGACGACGAGCTTACAAAAACCACTTCAAGCAACCAAAAAAGATCTCAAC 452
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 841 CMCAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 782
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 453 TCACACCTCACTGCAAGAGAGTACACACGACGACGACGACGACGACGACGAC 512
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 781 CMCAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 722
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 513 CACGACCAACCAACATCACTACACTGCTCAACCAACCAACAC 558
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 721 TCMCMCMAMAGCMAMCMMTTMAACCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 676
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 7
CNS01224/c
LOCUS 1097 bp DNA linear GSS 26-JUL-1999
DEFINITION Drosophila melanogaster genome survey sequence sp6 end of BAC
BACN08M09 of DrosBAC library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL102202
VERSION AL102202.1 GI:5613813
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 1097)
Genoscope.

AUTHORS Direct Submission
TITLE BP 191 91006 EVRY cedex - FRANCE (E-mail : seget@genoscope.cns.fr
JOURNML - Web : www.genoscope.cns.fr)

COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the European Drosophila Genome Project (EDGP) -
<http://www.edgp.ebi.ac.uk> - This Drosophila melanogaster BAC
library (Dros BAC) was made by Alain Billand at CEPH (Centre
d'Etude du Polymorphisme Humain) with funding provided by a MRC

project grant. The DNA was prepared from embryos by Alain Bucheton
and Genevieve Payan. It has been constructed in the vector
pBelobAC11.

FEATURES
source
location/Qualifiers
1. 1097
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BACN08M09"
/clone_lib="DrosBAC"
/plasmid="pBelobAC11"
/note="end : sp6"

BASE COUNT 169 a 119 c 196 g 398 t 215 others
ORIGIN

Query Match 13.4%; Score 95.8; DB 17; Length 1097;
Best Local Similarity 32.7%; Pred. No. 1.3e-08;
Matches 199; Conservative 119; Mismatches 290; Indels 0; Gaps 0;

```
OY 13 CTACCACTGCAATCATACAGATGCAACAGCCAGATCAAGAACAAACCCCACTAC 72
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1095 MTMCMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 1036
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 73 CTACCTAGCATCCTCAGCTTGGATCAGCTTCTCCATCTGTCTGAATTTACATCA 132
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 1035 CMCAMCMMAACAMACACMCMCMAMACACACACACACACACACACACACACAC 976
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 133 ACACGACGACATAGTGTGATCAACACACGAGAGTCAAGTCAACCTGCACCCACACA 192
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 975 CMMACACACACACACACACACACACACACACACACACACACACACACACACAC 916
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 193 GTCAAGATTAATAACACACACACACACACACACACACACACACACACACAC 252
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 915 AMACACACACACACACACACACACACACACACACACACACACACACACACAC 856
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 253 CGCCAAACCAACACACACACACACACACACACACACACACACACACAC 312
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 855 AHAACAAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 796
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 313 GTACCCCTGACGATATGACGACCAACATCAACCTGCTGAGCTATCTGCAAAAGATAC 372
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 795 TTTTTTTTTTTTTACANCMACMCMCTYMCATTAATMMCMACACACATVACCCATCMH 736
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 373 AACCAAAAACAGGAAAGAAACACACACAGCTTCAAAAAACCAACCTTCAAGACA 432
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 735 WMTVACMCMYACMAMMAACAMMMAMAMAMATCMHAMAMACAAAMCMACACACACAA 676
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 433 ACCAAAAAATCTCAACCTCAACCTCAACCTCAACCAAGAAAGTACCACACACAC 492
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 675 AAAAMCMAMACACACCCCMCACACCCGCCGCCGCCGCCGCCGCCGCCGCCGCC 616
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 493 ACAGAGAGGCAACATCAACACACACCAAAACAAATCACTCACTGCTCACCAC 552
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 615 MAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHAHA 556
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 553 AACACACAGGAAATTCAAAATCTCAAGTCAAAATGGAACCTTCCACCTCACCCTCC 612
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 555 TACMCMACACACACACACACACACACACACACACACACACACACACACACAC 496
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
OY 613 GAAGGCAA 620
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 495 ACAACCMH 488
      |||:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
```

RESULT 8
CNS009KS
LOCUS 791 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR19F04 of RPCL-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION AL053801
VERSION AL053801.1 GI:4935176
KEYWORDS GSS.
SOURCE Drosophila melanogaster.


```

QY 659 CATCAACCCCTCATCTCCACCAACACA 687
| | | | | | | | | | | | | | | | |
Db 966 CACCAACAACCAAAACCCACCAACAAA 994

RESULT 13
CNS00600/c 884 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BAC14N21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL065923
VERSION AL065923.1 GI:4944891
KEYWORDS GSS.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 884)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage;
BP 191 91006 Evry cedex - FRANCE (E-mail: sequef@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..884
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BAC14N21"
/clone_lib="RPCI-98"
/note="end : T7"

BASE COUNT 230 a 62 c 139 g 124 t 329 others
ORIGIN
Query Match 13.1%; Score 93.6; DB 17; Length 884;
Best Local Similarity 17.4%; Pred. No. 3.4e-08;
Matches 57; Conservative 167; Mismatches 103; Indels 0; Gaps 0;

QY 369 ACCAACAACCAAGAGGAAAGAAACACCAACCAAGCTCAAAAAACCACTTCAA 428
| | | | | | | | | | | | | | | | |
Db 884 MAMMMMMVMMHTKKKKTTTHMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 825

QY 429 GACAACCAAAAAAGATCTCAAACTCAAAACCTAAACCAAGGAGTACCCACCA 488
| | | | | | | | | | | | | | | | |
Db 824 AAMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 765

QY 489 GCCCAACAAGACCAACCATCAACACCAACCAACCAACATCACTAGCTGTAC 548
| | | | | | | | | | | | | | | | |
Db 764 MAMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 705

QY 549 CAACAACACCAAGAGTCAAAAACCTCAAGTCAATGGAACCTTCCACTCAACCTC 608
| | | | | | | | | | | | | | | | |
Db 704 CSMMAACMCACMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 645

QY 609 CTCGAAGGCAATCTTAAGCCCTTCTCAAGTCTCCAAACATCCAGACCACTCAAC 668
| | | | | | | | | | | | | | | | |

```

```

Db 644 MCCCCCCCMMMAACAACACAAAMAMAMAMAMAMAMAMAMAMAMAMAMAMAM 585
QY 669 CTCATCTCCACCCCAACACACACGCGCA 695
| | | | | | | | | | | | | | | | |
Db 584 MCCAMCTCACCAAMMCAAHMMAAAMMA 558

RESULT 14
A0743309 969 bp DNA linear GSS 16-JUL-1999
LOCUS HS_5387_B2_B07_SP6 RPCI-11 Human Male BAC Library Homo sapiens
DEFINITION genomic clone Plate=963 Col=14 Row=D, DNA sequence.
ACCESSION A0743309
VERSION A0743309.1 GI:5520831
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 969)
AUTHORS Mahalinas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
COMMENT Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome.
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
99380589
Contact: Mahalinas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallaceu.washington.edu
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (Inforesgen.com). BAC end web server:
http://www.hnsc.washington.edu
Plate: 963 row: D column: 14
Seq primer: SP6
Class: BAC ends
High quality sequence stop: 969.

FEATURES
source
1..969
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=963 Col=14 Row=D"
/clone_lib="RPCI-11 Human Male BAC Library"
/sex="male"
/note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
Male blood DNA was isolated from one randomly chosen donor
and partially digested with a combination of EcoRI and
EcoRI Methylase. Size selected DNA was cloned into the
pBACe3.6 vector at EcoRI sites"

BASE COUNT 647 a 294 c 7 g 5 t 16 others
ORIGIN
Query Match 13.1%; Score 93.6; DB 17; Length 969;
Best Local Similarity 48.0%; Pred. No. 3.4e-08;
Matches 286; Conservative 0; Mismatches 307; Indels 3; Gaps 1;

QY 119 AAATTACATCAACAACCAACCATCTAGCTTCAACACACAGGAGTCAAGTCAAC 178
| | | | | | | | | | | | | | | | |
Db 56 AACTTAACAACAACTCAANAAGCAAAACAAAMNNAACAACAACAAACCAACCA 115

QY 179 TGCACCCACACAGTCAAGACTAAACCAACAACCAACCAACCAACCAACG 238
| | | | | | | | | | | | | | | | |
Db 116 CAACAACCAACAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 175

QY 239 CCACTCAAAACCAACCAACCAACCAACCAACCAACCAATATGATTTTCACTTGG 298
| | | | | | | | | | | | | | | | |

```

Page 10

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="Plate=963 Col=4 Row=D"

```

Search completed: May 1, 2003, 12:33:32
Job time : 890.239 secs

PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis

PS Claim 9; Fig 3; 67pp; English.

XX The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternative
CC signal and expression sequences, for example the chimeric G protein
CC activated may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G proteins
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab.
CC by usual immunisation and cell fusion methods. This truncated G
CC protein is secreted since it lacks a transmembrane domain.

XX Sequence 232 AA:

Query Match 100.0%; Score 1229; DB 20; Length 232;
Best Local Similarity 100.0%; Pred. No. 3e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSQTTTILASTTPGVKSNL 60
DB 1 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSQTTTILASTTPGVKSNL 60
QY 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNKNDFHEVENFVPCISCSNNPTCMAIC 120
DB 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNKNDFHEVENFVPCISCSNNPTCMAIC 120
QY 121 KRIPNKKPGKTKTKTKTKTKTKTKDKLPQTTPKPEVPTTKPEEPTINTTKNTITTT 180
DB 121 KRIPNKKPGKTKTKTKTKTKTKTKDKLPQTTPKPEVPTTKPEEPTINTTKNTITTT 180
QY 181 LLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 232
DB 181 LLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 232

RESULT 2
AAR39286
ID AAR39286 standard; Protein: 298 AA.

AC AAR39286;

DT 13-JAN-1994 (first entry)

XX Respiratory syncytial virus (RSV) G protein.

KW PIV; RSV; multimeric; hybrid; pathogen; chimeric protein; vaccine.

OS Respiratory syncytial virus.

PN WO9314207-A.

PD 22-JUL-1993.

PF 05-JAN-1993; 93WO-CA00001.

PR 06-JAN-1992; 92GB-0000117.

PA (CONN-) CONNAGHT LAB LTD;

PI Ewasyslyn ME, Klein MH;

DR WPI: 1993-243222/30.

XX N-PSDB; AAQ45686.

PT Multimeric hybrid genes and their chimeric proteins - are

PT vaccines against multiple pathogenic infections e.g.
PT para-influenza virus and respiratory syncytial virus

PS Claim 11; Figure 7A-7D; 80pp; English.

XX A novel multimeric hybrid gene is used as a vaccine. The gene
CC consists of two gene sequences which are linked and encode antigenic
CC regions, these two sequences being derived from two different
CC pathogens (para-influenza virus (PIV) and respiratory syncytial virus
CC (RSV)). The gene sequences that are particularly used are those
CC which encode PIV-3 F and HN proteins (AAQ45683, AAQ45684) and RSV F and
CC G proteins (AAQ45685, AAQ45686).

XX Sequence 298 AA:

Query Match 100.0%; Score 1229; DB 14; Length 298;
Best Local Similarity 100.0%; Pred. No. 4e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSQTTTILASTTPGVKSNL 60
DB 67 HKVLTTLAIQDATSQIKNTPTPTLYLQDPQLGISFSNLSEITSQTTTILASTTPGVKSNL 126
QY 61 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNKNDFHEVENFVPCISCSNNPTCMAIC 120
DB 127 OPTVTKTKNTTTTQTOPSKPTTKORONKPPNKNDFHEVENFVPCISCSNNPTCMAIC 186
QY 121 KRIPNKKPGKTKTKTKTKTKTKTKDKLPQTTPKPEVPTTKPEEPTINTTKNTITTT 180
DB 187 KRIPNKKPGKTKTKTKTKTKTKTKDKLPQTTPKPEVPTTKPEEPTINTTKNTITTT 246
QY 181 LLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 232
DB 247 LLTNNTGNPKLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPPTTRQ 298

RESULT 3
AAW96313
ID AAW96313 standard; Protein: 298 AA.

AC AAW96313;

DT 28-JUN-1999 (first entry)

XX Membrane bound G protein of respiratory syncytial virus.

KW G protein; respiratory syncytial virus; RSV; recombinant vector;

KW vaccine; immune response; immunogenicity; tPA; antibody;

XX tissue plasminogen activator.

OS Respiratory syncytial virus.

PN WO9904010-A1.

PD 28-JAN-1999.

PF 16-JUL-1998; 98WO-CA00697.

PR 18-JUL-1997; 97US-0896442.

PA (CONN-) CONNAGHT LAB LTD.

PI Klein MH, Li X, Sambhara S;

DR WPI: 1999-132254/11.

XX N-PSDB; AAX08421.

PT Immunogenic composition for generating antibodies against
PT respiratory syncytial virus - comprises non-replicating vector
PT containing the protein G sequence, useful in protective vaccines and
PT to raise antibodies for diagnosis

PS Claim 4; Fig 2; 67pp; English.

XX The respiratory syncytial virus (RSV) G protein can be used in
CC vaccines by inserting the G protein gene into a non-replicating
CC vector. The G protein is placed under the control of alternating
CC signal and expression sequences, for example the chimeric G protein
CC produced may also comprise the signal peptide of tissue plasminogen
CC activator (tPA). The recombinant vector may also comprise sequences
CC upstream of the G protein gene which enhance the G protein's
CC immunoprotective ability. The resulting immunogenic composition will
CC generate antibodies directed against the RSV G protein when
CC administered to a host organism. The composition is useful as a
CC vaccine to immunise against RSV-associated disease, particularly
CC resulting in a balanced Th1/Th2 immune response and for raising Ab,
CC by usual immunisation and cell fusion methods.

SO Sequence 298 AA:

Query Match 100.0%; Score 1229; DB 20; Length 298;
Best Local Similarity 100.0%; Pred. No. 4e-89;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDAPATSOIKNTPTLYLTQDPOLGISFNSLSITTSQTTTLASTPGVKSML 60
DB 67 HKVTLTALIQDAPATSOIKNTPTLYLTQDPOLGISFNSLSITTSQTTTLASTPGVKSML 126
QY 61 QPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEFVFNVPSCISGNNPTCWAIC 120
DB 127 QPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEFVFNVPSCISGNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPK 180
DB 187 KRIPNKKPGKKTTPK 246
QY 181 LTNNTTGNPKLTSOMETFHSTSEGNLSPQSVSTSEHPSQSPSPPTTRQ 232
DB 247 LTNNTTGNPKLTSOMETFHSTSEGNLSPQSVSTSEHPSQSPSPPTTRQ 298

RESULT 4
AAP70845

ID AAP70845 standard; protein; 298 AA.

AC AAP70845;

DT 05-APR-1991 (first entry)

DE Sequence of human respiratory syncytial virus (HRSV) A2 strain
DE G protein.

KW Vaccine.

OS Human respiratory syncytial virus (HRSV).

PN W08704185-A.

PD 16-JUL-1987.

PF 23-DEC-1986; 86MO-US02756.

PR 14-JAN-1986; 86US-0818740.

PA (UYNC-) UNIV OF N CAROLINA.

PA (WERTZ) WERTZ G W.

PA (WERTZ) WERTZ G W.

DR WPI; 1987-206300/29.

DR N-PSDB; AAN70784.

XX Vaccines for human respiratory virus - comprising proteins or
PT fragment encoded by a DNA sequence coding for human respiratory
PT syncytial virus proteins.

PS Disclosure; Chart 13; 57pp; English.

XX A novel plasmid which comprises a DNA sequence encoding this
CC protein, and the protein itself, are claimed, for use as HRSV
CC vaccines. The vaccine can be administered to pregnant women or to
CC women of child bearing age to stimulate maternal antibodies.
CC Infants can also be vaccinated at 2-3 months of age.

SO Sequence 298 AA:

Query Match 91.1%; Score 1120; DB 8; Length 298;
Best Local Similarity 92.7%; Pred. No. 1.6e-80;
Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

QY 1 HKVTLTALIQDAPATSOIKNTPTLYLTQDPOLGISFNSLSITTSQTTTLASTPGVKSML 60
DB 67 HKVTLTALIQDAPATSOIKNTPTLYLTQDPOLGISFNSLSITTSQTTTLASTPGVKSML 126
QY 61 QPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEFVFNVPSCISGNNPTCWAIC 120
DB 127 QPTTVKTKNTTTQTQPSKPTTKORONKPPNKNDFEFVFNVPSCISGNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPK 180
DB 187 KRIPNKKPGKKTTPK 246
QY 181 LTNNTTGNPKLTSOMETFHSTSEGNLSPQSVSTSEHPSQSPSPPTTRQ 232
DB 247 LTNNTTGNPKLTSOMETFHSTSEGNLSPQSVSTSEHPSQSPSPPTTRQ 298

RESULT 5
AAR25302

ID AAR25302 standard; protein; 298 AA.

AC AAR25302;

DT 03-MAR-1993 (first entry)

DE HSRV glycoprotein G (gpG).

KW Vaccine; human respiratory syncytial virus; HRSV; F; G; 22K; 9.5K;

KW major capsid protein; N.

OS Human respiratory syncytial virus strain A2.

PN US5149650-A.

PD 22-SEP-1992.

PF 14-JAN-1986; 86US-0818740.

PR 14-JAN-1986; 86US-0818740.

PR 13-JUL-1988; 88US-0218737.

PA (UYNC-) UNIV NORTH CAROLINA.

PA Collins PL, Wertz GW;

DR WPI; 1992-340247/41.

DR N-PSDB; AAO29623.

XX Vaccines for human respiratory virus - include structural genes
PT coding for native structural viral proteins and immunogenic
PT fragments

XX Disclosure; Page 18; 21pp; English.

XX The sequences of mRNA encoding HRSV structural proteins are given in
CC AA029622-26. The proteins are F, G, 22K, 9.5K and major capsid
CC protein N. The sequences and encoded proteins are useful for
CC preparing vaccines against HRSV. The vaccines can be used to confer
CC immunity against respiratory tract infections on human subjects.

Seq	Sequence	298 AA	Query Match	91.1%	Score 1120	DB 13	Length 298
	Best Local Similarity	92.7%	Pred. No. 1,66-80				
	Matches 215	Conservative	5	Mismatches 12	Indels	0	Gaps 0
QY	1	HKVLTPTAIQDTSQIKNTPTPTVLYNQDPLGISPSNLSEITQTTIILASTTPGVKSNL 60					
Db	67	HKVLTPTAIQDTSQIKNTPTPTVLYNQDPLGISPSNLSEITQTTIILASTTPGVKSTL 126					
QY	61	QPTVTKTKNTTQTQPSKPTTKQKQNPKNPNDDHFEVFNNVPCSSCSNNPTCAIC 120					
Db	127	QSTVTKTKNTTQTQPSKPTTKQKQNPKNPNDDHFEVFNNVPCSSCSNNPTCAIC 186					
QY	121	KRIPIPNKPKSKKTTTKTKKPTFKTKRKDLKPQTKPEVPTKPEPTINTKTNIITTT 180					
Db	187	KRIPIPNKPKSKKTTTKTKKPTFKTKRKDLKPQTKPEVPTKPEPTINTKTNIITTT 246					
QY	181	LITNNITGNKFLTSQMEFHTSSSEGNLSPSQVSTTSSHPQSSSPENITRQ 232					
Db	247	LITNNITGNELTSQMEFHTSSSEGNLSPQVSTTSSHPQSSSPENITRQ 298					

	RESULT	6
AAW47605	ID	AAW47605 standard; Protein; 298 AA.
XX	AC	AAW47605;
XX	DT	11-JUN-1998 (first entry)
XX	DE	HRSV glycoprotein G.
XX	KM	HRSV; glycoprotein F; gpF; glycoprotein G; gpG; vaccine.
XX	OS	Human respiratory syncytial virus.
XX	PN	U55716823-A.
XX	PD	10-FEB-1998.
XX	PF	12-MAY-1997; 97US-0854783.
XX	PR	13-JUL-1988; 88US-0218737.
XX	PR	14-JAN-1986; 86US-0818740.
XX	PR	23-DEC-1986; 86MO-US02756.
XX	PR	11-JUN-1992; 92US-0891171.
XX	PR	12-MAY-1997; 97US-0854783.
PA	PA	(PHAA) PHARMACIA & UPJOHN CO.
PI	PI	Collins PL, Wertz GW;
XX	DR	WPI; 1998-144802/13.
XX	DR	N-PDB; AAV18736.
PT	PT	Production of human respiratory syncytial virus glyco-protein F or G
PT	PT	- by culturing eukaryotic host cells transfected with corresponding
XX	PT	DNA
PS	PS	Example 1; Columns 27-28; 17pp; English.
CC	CC	The present sequence was used in the development of a novel method
CC	CC	for the production of human respiratory syncytial virus (HRSV)
CC	CC	glycoprotein F (gpF) or glycoprotein G (gpG). The method comprises
CC	CC	culturing eukaryotic host cells transfected with an isolated DNA
CC	CC	sequence encoding HRSV gpF or gpG. The gp can be used to prepare
CC	CC	vaccines against HRSV.
SQ	SQ	Sequence 298 AA:
	Query Match	91.1%; Score 1120; DB 19; Length 298;
	Best Local Similarity	92.7%; Pred. No. 1,6e-80;

	Matches	215;	Conservative	5;	Mismatches	12;	Indels	0;	Gaps	0;
QY	1	HRYVLTATAIQDATSQIKNTPTPVYLQDPQLGCSFNSNLSEITSCOTIIILLASTPGVKSNL	60							
Db	67	HKVPPTAIIAQDPAISQIKNTPEPTFLQNQLDISPSNSENSEISQIITIIILASTPGVKSL	126							
QY	61	QPTTVTKKNTTQQPQSPKEPTKKOKONKPNNKDHEFEVFNFVPCISNNPTCMAIC	120							
Db	127	QSTTVKKNNTTTQTQPSPKPTTKQKONKPPSKPNNDHFEFEVFNVPVICISNNPTCMAIC	186							
QY	121	KRIPIKKPKGKKTTTKPKPKTEFTTKKLDKLPQTMKPKVEPVPTTKPEEPTINTKTNIITT	180							
Db	187	KRIPIKKPKGKKTTTKPKPKTEFTTKKLDKLPQTMKPKSVPTTKPEEPTINTKTNIITT	246							
QY	181	LLEINNTGNPKLTLSQMETFHSSTSSEGNLSPSOVSSTTSSEHPSSPNTTRQ	232							
Db	247	LITSNTGNBELTQSOMETFHSTSSSEGNSPSGVSTTSSEYPQSPSPNPTRQ	298							
RESULT 7										
AAU74676										
XX	AC	AAU74676 standard; Protein; 298 AA.								
XX	AC	AAU74676;								
XX	-DT	09-APR-2002 (first entry)								
DE	XX	Respiratory syncytial virus G protein.								
XX	XX									
XX	RSV; G protein; heavily glycosylated protein; anti-naemic; antiviral;									
XX	vaccine; gene therapy; parainfluenza; Sendai virus; PMV;									
XX	antiviral chemotherapeutic compound; humoral response;									
XX	cellular immune response; HIV; paediatric respiratory disease;									
XX	globulin gene transfer; sickle cell disease; beta-thalassemia;									
XX	human immunodeficiency virus infection; HIV.									
OS	Human respiratory syncytial virus.									
XX	WO200192548-A2.									
XX	06-DEC-2001.									
XX	22-MAY-2001; 2001WO-US1610.									
PR	01-JUN-2000; 2000US-208701P.									
XX	(STUD-) ST JUDE CHILDREN'S RES HOSPITAL.									
XX	Portner A, Takimoto T;									
DR	WPI; 2002-130534/17.									
DR	N-PDB; AAS21045.									
PT	Recombinant Sendai virus useful in vaccines to protect infection by									
PT	parainfluenza, comprises exogenous nucleic acid encoding									
XX	parainfluenza protein or its antigenic fragment -									
XX	Disclosure; Page 48; 57pp; English.									
CC	The invention relates to a recombinant Sendai virus comprising an									
CC	exogenous nucleic acid encoding a parainfluenza (PMV) protein or its									
CC	antigenic fragment. The virus may be administered in combination									
CC	with an antiviral chemotherapeutic compound. Two or more viruses									
CC	expressing different PMV proteins may be co-administered. Compositions									
CC	comprising the virus are useful for eliciting a humoral and/or									
CC	cellular immune response to a PMV in a mammal, particularly a human.									
CC	Further a recombinant Sendai virus comprising an exogenous nucleic acid									
CC	encoding a second PMV protein is also administered and priming and/or									
CC	boosting humoral or cellular immune response comprises administering									
CC	one or more of a recombinant or isolated PMV protein or its antigenic									
CC	fragment, a DNA vaccine encoding the same, and a non-Sendai viral									
CC	vector encoding a PMV protein. The recombinant virus is useful as an									
CC	effective vaccine against HIV or RSV (the major causes of paediatric									

CC	vector encoding a PMV protein. The recombinant virus is useful as an effective vaccine against HIV or RSV (the major causes of paediatric
CC	fragment, a DNA vaccine encoding the same, and a non-Sendai viral
CC	one or more of a recombinant or isolated PMV protein or its antigenic
CC	boosting humoral or cellular immune response comprises administering
CC	encoding a second PMV protein is also administered and priming and/or
CC	further a recombinant Sendai virus comprising an exogenous nucleic acid
CC	cellular immune response to a PMV in a mammal, particularly a human,
CC	comprising the virus are useful for eliciting a humoral and/or
CC	expressing different PMV proteins may be co-administered. Compositions
CC	with an antiviral chemotherapeutic compound. Two or more viruses
CC	antigenic fragment. The virus may be administered in combination
CC	exogenous nucleic acid encoding a paramyxovirus (PMV) protein or its
XX	The invention relates to a recombinant Sendai virus comprising an
PS	disclosure; Page 48; 57pp; English.
XX	
XX	
PT	parainfluenza virus protein or its antigenic fragment - encoding
PT	parainfluenza virus protein or its antigenic fragment - encoding
DR	N-PSDB; AAS21045.
XX	
XX	
PA	(SUDJ-) ST JUDE CHILDREN'S RES HOSPITAL.
PI	Portner A, Takimoto T;
XX	
XX	
XX	WPI; 2002-130534/17.
XX	
XX	
PF	22-MAY-2001; 2001WO-US16610.
XX	
PR	01-JUN-2000; 2000US-208701P.
XX	
OS	Human respiratory syncytial virus.
PN	WO200192548-A2.
XX	
PD	06-DEC-2001.
XX	
XX	
KM	cellular immune response; hPIV; paediatric respiratory disease;
KM	globulin gene transfer; sickle cell disease; beta-thalassemia;
KM	human immunodeficiency virus infection; HIV.
XX	
DE	Respiratory syncytial virus G protein.
XX	
XX	
KW	RSV; G protein; heavily glycosylated protein; anti-nausea; antiviral;
KW	vaccine; gene therapy; paramyxovirus; sendai virus; PMV;
KW	antiviral chemotherapeutic compound; humoral response;
ID	AAU74676 standard; Protein; 298 AA.
AC	
XX	
DT	09-APR-2002 (first entry)
XX	
XX	
RESULT 7	
AAU74676	

CC respiratory disease) and also to express any gene of
 CC interest in target cells, providing a positive medical impact on
 CC impaired cells. Wild-type globin gene transfer (i.e. gene therapy)
 CC into stem cells effects a cure for sickle cell disease or beta-
 CC thalassemia. The recombinant virus may also prove effective in
 CC conferring immunity to human immunodeficiency virus (HIV) infection.
 CC The Sendai virus replicates at level that is high enough to
 CC induce sufficient immunity, but does not cause any harm to human
 CC recipient. The present sequence represents a respiratory syncytial
 CC virus (RSV) G protein (heavily glycosylated protein), a PMV protein
 CC suitable for expression by the recombinant virus of the invention.

XX Sequence 298 AA:

Query Match 91.1%; Score 1120; DB 23; Length 298;
 Best Local Similarity 92.7%; Pred. No. 1.6e-80;

Matches 215; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

OY 1 HKVTLTAIIODATSOIKNTTPYLTODPOLGISFNSLSEITSOITLILASTPGVKSNL 60
 DB 67 HKVPTTAIIODATSOIKNTTPYLTQNPOLGISPSNPSEITSOITLILASTPGVKS 126
 OY 61 OPTVTKNTTTOTOPSKPTTKORONKPPNKPNNDPHEFVFNVPSCISNNPTCAIC 120
 DB 127 OSTVTKNTTTOTOPSKPTTKORONKPPNKPNNDPHEFVFNVPSCISNNPTCAIC 186
 OY 121 KRIPNKKPGKTTTKPKPKFTTKKDLKPQTTPKKEVPTTKPTEPTINTKTNTTT 180
 DB 187 KRIPNKKPGKTTTKPKPKFTTKKDLKPQTTPKKEVPTTKPTEPTINTKTNTTT 246
 OY 181 LITNTTGNKRLTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPNTTRQ 232
 DB 247 LITSNTGNPELTSQMETFHSSTSEGNLSPSOVSTTSEHPSQSSPNTTRQ 298

RESULT 8

AAP90441
 ID AAP90441 standard; protein; 681 AA.

XX AAP90441:

XX 01-NOV-1989 (first entry)

DE Chimeric human respiratory syncytial virus glycoproteins F and G.

XX Chimeric polypeptide; human respiratory syncytial virus;

KW protein F; protein G; vaccine.

XX Human respiratory syncytial virus.

XX WO8905823-A.

XX 29-JUN-1989.

XX 31-OCT-1988; 88WO-US03784.

XX 23-DEC-1987; 87US-0137387.

XX (UPJO) UPJOHN CO.

XX Wathen M;

XX WPI; 1989-206593/28.

PT Chimeric human respiratory syncytial virus polypeptides(s)

PT - contg. immunogenic fragments from RSV glycoproteins

CC F and G, for vaccine prodn.

XX Claim 3; page 47-48; 50pp; English.

CC Chimeric polypeptide contg. a signal sequence and one or more

CC immunogenic fragments from both human respiratory syncytial virus

CC glycoproteins F and G. Can be used in vaccines. Hosts are, eg

CC E. coli, Chinese hamster ovary cells, murine C127 cells and

CC S. frugiperda.

XX Sequence 681 AA:

Query Match 73.3%; Score 901; DB 10; Length 681;
 Best Local Similarity 91.0%; Pred. No. 7.3e-63;

Matches 172; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 30 QLGISFNSLSEITSOITLILASTPGVKSNDLPOTVTKNTTTOTOPSKPTTKORONK 89
 DB 490 QLGISPSNPSEITSOITLILASTPGVKSNDLPOTVTKNTTTOTOPSKPTTKORONK 549
 OY 90 PNKPNNDPHEFVFNVPSCISNNPTCAICKRIPNKKPGKTTTKPKPKFTTKKDL 149
 DB 550 PSKPNNDPHEFVFNVPSCISNNPTCAICKRIPNKKPGKTTTKPKPKFTTKKDL 609
 OY 150 KPOTTPKKEVPTTKPTEPTINTKTNTTTLITNTTGNPKLTSQMETFHSSTSEGNLS 209
 DB 610 KPOTTPKKEVPTTKPTEPTINTKTNTTTLITNTTGNPKLTSQMETFHSSTSEGNLS 669
 OY 210 PSQVSTTSE 218
 DB 670 PSQVNSTSQ 678

RESULT 9

AAB67771
 ID AAB67771 standard; protein; 452 AA.

XX AAB67771:

XX 11-JUN-2001 (first entry)

DE Amino acid sequence of a fusion protein of P40 and RSV antigen.

XX Outer membrane protein A; OmpA; P40; enterobacteria; nasal composition;

KW RSV; RSV infection; lung; respiratory tract; vaccine.

XX Synthetic.

OS Klebsiella pneumoniae.

OS Respiratory syncytial virus.

XX WO200121203-A1.

XX 29-MAR-2001.

XX 22-SEP-2000; 2000WO-FR02626.

XX 23-SEP-1999; 99FR-0011888.

XX (FABR) FABRE MEDICAMENT SA PIERRE.

XX Corvalan N, Goestch L;

XX WPI; 2001-257929/26.

XX N-PSDB; AAF80153.

PT Vaccine against respiratory syncytial virus, comprises enterobacterial

PT outer membrane protein and viral immunogen, provides protective

XX response throughout the respiratory tract

XX Example 2; Page 31-32; 39pp; French.

CC The present sequence represents a fusion protein comprising a Klebsiella

CC pneumoniae outer membrane protein A (OmpA) designated P40 and a

CC respiratory syncytial virus (RSV) antigen. Enterobacterium OmpA proteins,

CC associated with an immunogenic peptide from RSV are used to prepare

CC a nasal composition that induces a protective response, against RSV

CC infection in the upper and lower (lung) respiratory tract. OmpA

CC potentiates the immune response to some immunogenic peptides,

CC eliminating the need for adjuvants. The method is useful for producing

CC vaccines for prevention or treatment of RSV infections.

XX Sequence 452 AA:
 SQ Query Match 46.6%; Score 572.5; DB 22; Length 452;
 Best Local Similarity 76.1%; Pred. No. 3.5e-37;
 Matches 108; Conservative 5; Mismatches 18; Indels 11; Gaps 2;

OY 25 LQDDPLGTSFSLSEITSOITTLASTTPGVKSNLOPTTKTKNTTTOTOPSPKPTTKO 84
 Db 322 LAPDRVEIEVKGKYEVTQ-----PGCG-----DPVTAKTKNTTTOTOPSPKPTTKO 370
 OY 85 RQKPPKPNNDHFENFVPCISNNPTGMAICKRIIPNKKPGKTTTKPKPTKPTKT 144
 Db 371 RQKPPKPNNDHFENFVPCISNNPTGMAICKRIIPNKKPGKTTTKPKPTKPTKT 430
 OY 145 TKKDLKPQTTKREVPPTTKPTE 166
 Db 431 TKDKRPQTTKREVPPTTKPVD 452

RESULT 10
 AAR95660
 ID AAR95660 standard; Protein: 349 AA.
 XX AAR95660;
 AC
 XX
 DT 21-JAN-1997 (first entry)
 XX
 DE Streptococcal protein G fragment BB fused to RSV hapten G2A.
 XX
 KW Streptococcus; protein G; carrier protein; immunogenicity; increase;
 KM enhance; vaccine; anti-viral; human serum albumin; binding;
 KW immunostimulation; respiratory syncytial virus; fusion protein.
 XX
 OS Chimeric Streptococcus sp.
 OS Chimeric Respiratory Syncytial Virus.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..23
 FT /label= signal_peptide
 FT Protein 24..242
 FT /label= BB
 FT Protein 247..347
 FT /label= G2A
 FT /note= "residues 130-230 of RSV"
 XX
 FT
 FT
 XX
 PN WO9614416-A1.
 PD 17-MAY-1996.
 XX
 PF 07-NOV-1995; 95WO-FR01466.
 XX
 PR 07-NOV-1994; 94FR-0013310.
 XX
 PA (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 PI Andreoni C, Binz H, Nguyen Ngoc T, Nygren PA, Stahl S;
 PI Uhlen M;
 XX
 DR MPI: 1996-251766/25.
 DR N-PSDB; AAT31647.
 XX
 PT Enhancing immunogenicity by coupling immunogen to serum
 PT albumin-binding protein - useful for preparing improved vaccines,
 PT e.g. against Respiratory Syncytial Virus
 XX
 PS Disclosure; Page 81-83; 102pp; French.
 XX
 CC The present sequence comprises a 219 amino acid fragment of
 CC Streptococcal protein G, designated BB, fused to an immunogen,
 CC designated G2A, derived from amino acids 130-230 of protein G from
 CC Respiratory Syncytial Virus sub-group A or B. The BB fragment is
 CC able to bind to human serum albumin and thereby enhance immunogenicity

CC of any antigen, hapten or immunogen that is covalently coupled to it.
 CC In this specific example, the BB fragment was found to induce T helper
 CC memory cells, leading to production of anti-G2A antibodies by stimulated
 CC B cells.
 XX
 SQ Sequence 349 AA:
 OY Query Match 45.7%; Score 562; DB 17; Length 349;
 Best Local Similarity 78.1%; Pred. No. 1.8e-36;
 Matches 107; Conservative 10; Mismatches 10; Indels 10; Gaps 3;

OY 37 NLSEITSOITTT-ILASTT-----PGVKSNOPTTKTKNTTTOTOPSPKPTTKOROK 89
 Db 216 NGKTLKGETTEAVADAATARSFNPILNSW--TVTKNTTTOTOPSPKPTTKOROK 272
 OY 90 PNKPNDHFENFVPCISNNPTGMAICKRIIPNKKPGKTTTKPKPTKPTKTKDL 149
 Db 273 PNKPNDHFENFVPCISNNPTGMAICKRIIPNKKPGKTTTKPKPTKPTKTKDKH 332
 OY 150 KPQTTKREVPPTTKPTE 166
 Db 333 KPQTTKREVPPTTKPVD 349

RESULT 11
 AAB68028
 ID AAB68028 standard; Protein: 349 AA.
 XX AAB68028;
 AC
 XX
 DT 29-JUN-2001 (first entry)
 XX
 DE Amino acid sequence of fusion protein comprising 2 G protein fragments.
 XX
 KW Aliphatic ammonium salt; immunogen; antigen; syncytial virus infection;
 KM G protein; fusion protein.
 XX
 OS Synthetic.
 OS Streptococcus sp.
 OS Respiratory syncytial virus.
 XX
 FH Key Location/Qualifiers
 FT Protein 120..230
 FT /note= "G protein fragment of VRS",
 XX
 XX
 PN FR2798292-A1.
 PD 16-MAR-2001.
 XX
 PF 09-SEP-1999; 99FR-0011284.
 XX
 PR 09-SEP-1999; 99FR-0011284.
 XX
 PA (FABR) FABRE MEDICAMENT SA PIERRE.
 XX
 PI Beck A, Klingner C, Nguyen TN;
 DR MPI: 2001-267782/28.
 DR N-PSDB; AAF84711.
 XX
 PT use of quaternary aliphatic ammonium salt and immunogen or antigen to
 PT combat respiratory syncytial virus infections
 XX
 PS Claim 16; Page 22-23; 35pp; French.
 XX
 CC The specification describes a method which uses quaternary aliphatic
 CC ammonium salts together with an immunogen or antigen to treat syncytial
 CC virus infections. The combination of the salt with the antigen or
 CC immunogen improves immunogenicity and equilibrates the Th1/Th2 immune
 CC response. The method is used for the treatment of respiratory syncytial
 CC virus infections. The present sequence represents a fusion protein,
 CC comprising an albumin binding domain of the G protein of Streptococcus
 CC sp. fused to a G protein fragment of respiratory syncytial virus (VRS,

CC	INJECTIONS.
XX	
50	Sequence 101 AA;

Best Local Similarity 99.0%; Pred. No. 8,8e-37;
Matches 100; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 TVKTNNTTTOTOPSKPTTKORONKPPKNNDFHEFVFNVPSCISNNPTCAICKRI 60
|||||

OY 124 PNKKPGKTTTKPTKKPTFKTTKDLKPQTTRKREVPPTKP 164
|||||
DB 61 PNKKPGKTTTKPTKKPTFKTTKDKHKPQTTRKREVPPTKP 101
|||||

RESULT 14

AAR95616 standard; Protein; 101 AA.

AAR95616;

10-FEB-1997 (first entry)

RSV sub-group A wild type protein G residues 130-230.

Heterologous protein; cell surface; Staphylococcus xyloso; protein G;
Staphylococcus carnosus; respiratory syncytial virus; wild type; mutant;
expression plasmid; fusion protein; streptococcal.

Human respiratory syncytial virus.

WO9614418-A1.

17-MAY-1996.

07-NOV-1995; 95WO-FR01465.

07-NOV-1994; 94FR-0013309.

(FABR) FABRE MEDICAMENT SA PIERRE.

Binz H, Nguyen Ngoc T, Nygren PA, Stahl S, Uhlen M;

WPI: 1996-251768/25.

N-PSDB; AAT27073.

Expression of respiratory syncytial virus protein G at cell surface
- of bacteria non-pathogenic for mammals, useful in orally active
vaccines

Claim 12; Page 13-14; 38pp; French.

The invention relates to the prodn. of a heterologous protein on the
surface of a cell, pref. Staphylococcus xyloso or S. carnosus, contg. a
fragment of the respiratory syncytial virus comprising residues 130-230
of the RSV protein G. This sequence is though to be the wild type
residues 130-230 from the RSV subgp. A. The heterologous protein may
also contain the subgp. B sequence (AAT27074). These sequences may be
mutated to replace the Cys residues at pos. 173 and 186 by residues
unable to form a disulphide bridge esp. Ser residues (see AAT27075-6).
Other mutations include substitution of the Phe residues at pos. 163,
165, 168 and/or 170 by Ser residues. The sequences encoding the wild
type or mutant RSV residues 130-230 are inserted into plasmid
PSE/mpi18BXM to create plasmid PSE/G2BXM (wild type sequence) or
PSE/G2subBXM (mutant sequence). The sequence is produced from PSE/BXM
as a fusion protein where the fusion is with the streptococcal protein G
serum albumin binding domain. These plasmids are introduced into
S. xyloso or S. carnosus for expression of the wild type or mutant
protein.

Sequence 101 AA;

Query Match 45.4%; Score 558; DB 17; Length 101;

Best Local Similarity 99.0%; Pred. No. 8.8e-37; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 1;

OY 64 TVKTKNTTTOTOPSKPTTKORONKPPKNNDFHEFVFNVPSCISNNPTCAICKRI 123
|||||
DB 1 TVKTKNTTTOTOPSKPTTKORONKPPKNNDFHEFVFNVPSCISNNPTCAICKRI 60
|||||

OY 124 PNKKPGKTTTKPTKKPTFKTTKDLKPQTTRKREVPPTKP 164
|||||
DB 61 PNKKPGKTTTKPTKKPTFKTTKDKHKPQTTRKREVPPTKP 101
|||||

RESULT 15

AAR97050 standard; Protein; 101 AA.

AAR97050;

20-JAN-1997 (first entry)

Respiratory syncytial virus antigenic fragment G2A.

Streptococcus; protein G; carrier protein; immunogenicity; increase;
enhance; vaccine; anti-viral; human serum albumin; binding;
immunostimulation; RSV; hapten; antigen.

Respiratory Syncytial Virus.

WO9614416-A1.

17-MAY-1996.

07-NOV-1995; 95WO-FR01466.

07-NOV-1994; 94FR-0013310.

(FABR) FABRE MEDICAMENT SA PIERRE.

Andreoni C, Binz H, Nguyen Ngoc T, Nygren PA, Stahl S;

Uhlen M;

WPI: 1996-251766/25.

N-PSDB; AAT31649.

Enhancing immunogenicity by coupling immunogen to serum
albumin-binding protein - useful for preparing improved vaccines,
e.g. against Respiratory Syncytial Virus

Claim 21; Page 29-30; 102pp; French.

The present sequence is that of an immunogen (designated G2A) derived
from amino acids 130-230 of protein G from Respiratory Syncytial Virus
sub-group A or B. When fused to a 219 amino acid fragment of G2A
Streptococcal protein G, designated BB, the immunogenicity of G2A
was enhanced. Specifically, the BB fragment was found to induce T
helper memory cells, leading to production of anti-G2A antibodies by
stimulated B cells. The BB fragment is able to bind to human serum
albumin and thereby enhance immunogenicity of any antigen, hapten or
immunogen that is covalently coupled to it.

Sequence 101 AA;

Query Match 45.4%; Score 558; DB 17; Length 101;

Best Local Similarity 99.0%; Pred. No. 8.8e-37; Indels 0; Gaps 0;

Matches 100; Conservative 0; Mismatches 1;

OY 64 TVKTKNTTTOTOPSKPTTKORONKPPKNNDFHEFVFNVPSCISNNPTCAICKRI 123
|||||
DB 1 TVKTKNTTTOTOPSKPTTKORONKPPKNNDFHEFVFNVPSCISNNPTCAICKRI 60
|||||

OY 124 PNKKPGKTTTKPTKKPTFKTTKDLKPQTTRKREVPPTKP 164
|||||
DB 61 PNKKPGKTTTKPTKKPTFKTTKDKHKPQTTRKREVPPTKP 101
|||||

Search completed: May 1, 2003, 13:44:22
Job time : 29.3283 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:43:28 ; Search time 12.2566 Seconds
(without alignments)
556.934 Million cell updates/sec

Title: US-09-462-816-4

Sequence: 1 HKVLTITAIIDATSIQIKNT.....VSTTSEHPSPSPPTTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued_Patents_AA.*
- 2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*
- 3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*
- 4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*
- 5: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*
- 6: /cgn2_6/ptodata/1/1aa/PCtus.COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1229	100.0	298	2	US-08-467-963C-8
2	1229	100.0	298	2	US-08-838-189D-8
3	1229	100.0	298	3	US-08-852-344D-8
4	1229	100.0	298	3	US-08-344-639E-8
5	1229	100.0	298	4	US-08-467-969A-8
6	1229	100.0	298	4	US-08-467-961A-8
7	1229	100.0	298	4	US-08-001-554A-8
8	901	73.3	681	6	5194595-19
9	222	18.1	37	3	US-08-793-792-12
10	193	15.7	32	3	US-08-793-792-8
11	188	15.3	216	3	US-08-928-361B-8
12	188	15.3	1837	3	US-08-928-361B-5
13	187.5	15.3	1721	3	US-08-700-651-5
14	187.5	15.3	1721	3	US-08-928-361B-6
15	181.5	14.8	216	3	US-08-928-361B-27
16	174	14.2	28	3	US-08-793-792-4
17	173.5	14.1	249	3	US-08-700-651-15
18	170.5	14.1	249	3	US-08-928-361B-20
19	170.5	13.9	263	5	PCT-US91-08177-13
20	164.5	13.4	2476	4	US-08-276-967-2
21	162	13.2	941	4	US-07-757-022B-14
22	162	13.2	1022	4	US-07-757-022B-84
23	162	13.2	1038	4	US-07-757-022B-74
24	162	13.2	1049	4	US-07-757-022B-58
25	162	13.2	1140	4	US-07-757-022B-104
26	162	13.2	1270	4	US-07-757-022B-44
27	162	13.2	1311	4	US-07-757-022B-42

28	162	13.2	1313	4	US-07-757-022B-142	Sequence 142, App
29	162	13.2	1314	4	US-07-757-022B-50	Sequence 50, Appl
30	162	13.2	1320	4	US-07-757-022B-46	Sequence 46, Appl
31	162	13.2	1320	4	US-07-757-022B-60	Sequence 60, Appl
32	162	13.2	1354	4	US-07-757-022B-48	Sequence 48, Appl
33	162	13.2	1361	4	US-07-757-022B-40	Sequence 40, Appl
34	162	13.2	1363	4	US-07-757-022B-52	Sequence 52, Appl
35	162	13.2	1404	4	US-07-757-022B-2	Sequence 2, Appl1
36	162	13.2	1404	4	US-07-757-022B-62	Sequence 62, Appl
37	158	12.9	175	3	US-08-700-651-12	Sequence 12, Appl
38	158	12.9	175	3	US-08-928-361B-17	Sequence 17, Appl
39	153.5	12.5	878	4	US-09-556-706B-2	Sequence 2, Appl1
40	151	12.3	37	3	US-08-793-792-11	Sequence 11, Appl
41	150	12.2	30	4	US-08-836-504A-6	Sequence 6, Appl1
42	148	12.0	26	4	US-08-750-624-9	Sequence 9, Appl1
43	148	12.0	629	4	US-09-241-581B-6	Sequence 6, Appl1
44	148	12.0	629	4	US-08-265-428-6	Sequence 6, Appl1
45	148	12.0	629	5	PCT-US95-07721-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-08-467-963C-8
Sequence 8, Application US/08467963C

Patent No. 5968776

GENERAL INFORMATION:

APPLICANT: KLEIN, Michel H

APPLICANT: DU, Run-Pan

TITLE OF INVENTION: EMASYSBYN, Mary E

TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A CHIMERIC PROTEIN WHICH CONFERNS PROTECTION AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY STRUCTIAL VIRUS

TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY STRUCTIAL VIRUS

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,963C

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/838,189

FILING DATE: 16-Apr-1997

APPLICATION NUMBER: US 08/001,554

FILING DATE: 06-JAN-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: GB 9200117.1

FILING DATE: 06-JAN-1992

ATTORNEY/AGENT INFORMATION:

NAME: STEWART, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-474 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 298 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-467-963C-8

Query Match	100.0%	Score 1229	DB 2	Length 298
Best Local Similarity	100.0%	Pred. No. 2	4e-105	
Matches 232	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	HKVLTLLTAIIODATSOIKNTPTPTVLTODPQIGISFSNLSSEITSSQTTTIIASTPTGVKSNL	60	
Db	67	HKVLTLLTAIIODATSOIKNTPTPTVLTODPOLGISFSNLSSEITSSQTTTIIASTPTGVKSNL	126	
QY	61	OPTPTKANTTTTTOPOSPKPTKORONKPPNPNNDHFHEVNFAPCSISNNPTCMAIC	120	
Db	127	OPTPTKIKNTTTTTOPOSPKPTKORONKPPNPNNDHFHEVNFAPCSISNNPTCMAIC	186	
QY	121	KRIPKKPGKKTTPTKPKPKPTKTKDKLPQTKPKEVPTKPIEPTINTTKTINITYT	180	
Db	187	KRIPKKPGKKTTPTKPKPKPTKTKDKLPQTKPKEVPTKPIEPTINTTKTINITYT	246	
QY	181	LTNTNTGPKLTLSOMETPHSTSSSGNLSPSQVSTTSEHPSSPSPNTTRQ	232	
Db	247	LTNTNTGPKLTLSOMETPHSTSSSGNLSPSQVSTTSEHPSSPSPNTTRQ	298	

RESULT 2
 US-08-838-189D-8
 : Sequence 8, Application US/08838189D
 : Patent No. 598169
 : GENERAL INFORMATION:
 : APPLICANT: KLEIN, Michel H
 : APPLICANT: DU, Run-Pan
 : APPLICANT: EMASYSN, Mary E
 : TITLE OF INVENTION: MULTIMERIC HYBRID GENE ENCODING A
 : TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION AGAINST
 : TITLE OF INVENTION: PARAINFLUENZA VIRUS AND RESPIRATORY SYNCYTIAL VIRUS
 : NUMBER OF SEQUENCES: 38
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Sam & McBurney
 : STREET: 6th floor, 330 University Avenue
 : CITY: Toronto
 : STATE: Ontario
 : COUNTRY: Canada
 : ZIP: M5G 1R7
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: Patentln Release #1.0, Version #1.30
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/838,189D
 : FILING DATE: 16-APR-1997
 : CLASSIFICATION: 435
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US 08/001,554
 : FILING DATE: 06-JAN-1993
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: GB 9200117.1
 : FILING DATE: 06-JAN-1992
 : ATTORNEY/AGENT INFORMATION:
 : NAME: STEWART, Michael I
 : REGISTRATION NUMBER: 24,973
 : REFERENCE/DOCKET NUMBER: 1038-687 MTS:jb
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (416) 595-1155
 : TELEFAX: (416) 595-1163
 : INFORMATION FOR SEQ ID NO: 8:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 298 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : US-08-838-189D-8

Matches	2332	Conservative	0	Mismatches	0	Indels	0	Gaps	0
OY	1	HKVLTITAIIDDAISQIKNTPTTYLTOTOPOLGISNSLSETTSOTHTLASTPEGVKSNL	60						
Db	67	HKVLTITAIIDDAISQIKNTPTTYLTOTOPOLGISNSLSETTSOTHTLASTPEGVKSNL	126						
OY	61	OPTVTKNTNTTQTOPSPKPTTKORONPPNPKPNNDPHEEYFNFPVCSICSNNPCTMAIC	120						
Db	127	OPTVTKNTNTTQTOPSPKPTTKORONPPNPKPNNDPHEEYFNFPVCSICSNNPCTMAIC	186						
OY	121	KRIENKATGGKTTTPTKPPFTTKKOLKQJTYTKPREVPTTKPTEEPINTTKNITTT	180						
Db	187	KRIENKATGGKTTTPTKPPFTTKKOLKQJTYTKPREVPTTKPTEEPINTTKNITTT	246						
OY	181	LTNTNTNPKLTSSOMETFHSTSSGNLSPSSVSTTSEHPSPSSPMTTQ	212						
Db	247	LTNTNTNPKLTSSOMETFHSTSSGNLSPSSVSTTSEHPSPSSPMTTQ	298						

```

1      RESULT 3
2      US-08-852-344D-8
3      : Sequence 8, Application US/08852344D
4      : Patent No. 6017539
5      :
6      : GENERAL INFORMATION:
7      :
8      : APPLICANT: KLEIN, Michel H
9      :
10     : APPLICANT: DU, Run-Pan
11     :
12     : APPLICANT: EMASYSHYN, Mary E
13     :
14     : TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERS PROTECTION
15     :
16     : TITLE OF INVENTION: AGAINST PARAINFLUENZA VIRUS AND RESPIRATORY
17     :
18     : TITLE OF INVENTION: SYNCYTIAL VIRUS
19     :
20     : NUMBER OF SEQUENCES: 37
21     :
22     : CORRESPONDENCE ADDRESS:
23     :
24     : ADDRESSEE: Sim & McBurney
25     :
26     : STREET: 6th Floor, 330 University Avenue
27     :
28     : CITY: Toronto
29     :
30     : STATE: Ontario
31     :
32     : COUNTRY: Canada
33     :
34     : ZIP: M5G 1R7
35     :
36     : COMPUTER READABLE FORM:
37     :
38     : MEDIUM TYPE: Floppy disk
39     :
40     : COMPUTER: IBM PC compatible
41     :
42     : OPERATING SYSTEM: PC-DOS/MS-DOS
43     :
44     : SOFTWARE: Patent in Release #1.0, Version #1.30
45     :
46     : CURRENT APPLICATION DATA:
47     :
48     : APPLICATION NUMBER: US/08/852.344D
49     :
50     : FILING DATE: 07-MAY-1997
51     :
52     : CLASSIFICATION: 424
53     :
54     : PRIOR APPLICATION DATA:
55     :
56     : APPLICATION NUMBER: US 08/344.639
57     :
58     : FILING DATE: 14-NOV-1994
59     :
60     : PRIOR APPLICATION DATA:
61     :
62     : APPLICATION NUMBER: GB 9200117.1
63     :
64     : FILING DATE: 06-JAN-1992
65     :
66     : ATTORNEY/AGENT INFORMATION:
67     :
68     : NAME: STEWART, Michael I
69     :
70     : REGISTRATION NUMBER: 24,973
71     :
72     : REFERENCE/DOCKET NUMBER: 1038-688 MIS:jb
73     :
74     : TELECOMMUNICATION INFORMATION:
75     :
76     : TELEPHONE: (416) 595-1155
77     :
78     : TELEFAX: (416) 595-1163
79     :
80     : INFORMATION FOR SEQ ID NO: 8:
81     :
82     : SEQUENCE CHARACTERISTICS:
83     :
84     : LENGTH: 298 amino acids
85     :
86     : TYPE: amino acid
87     :
88     : STRANDEDNESS: single
89     :
90     : TOPOLOGY: linear
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :
622    :
623    :
624    :

```

|||||
Db 67 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 126
Qy 61 QPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 186
Qy 121 KRIPNKKPGKTKTTTKPKPTTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 180
Db 187 KRIPNKKPGKTKTTTKPKPTTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 246
Qy 181 LTNNNTGNPKLTQSOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTRQ 232
Db 247 LTNNNTGNPKLTQSOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTRQ 298

RESULT 4

US-08-344-639E-8
; Sequence 8, Application US/08344639E
; Patent No. 6033668
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasyszyn, Mary E
; TITLE OF INVENTION: CHIMERIC PROTEIN WHICH CONFERES
; TITLE OF INVENTION: PROTECTION AGAINST PARAINFLUENZA VIRUS
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,639E
; FILING DATE: 14-NOV-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-391 MIS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-344-639E-8

Query Match 100.0%; Score 1229; DB 3; Length 298;
Best Local Similarity 100.0%; Pred. No. 2,4e-105;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 60
Db 67 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 126
Qy 61 QPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKNPNNDPFHEVFNFVPCISNNPTCAIC 186
Qy 121 KRIPNKKPGKTKTTTKPKPTTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 180
Db 187 KRIPNKKPGKTKTTTKPKPTTKTKDKLPQTTKPKKEVPTTKPTEPTINTTKNTTTT 246
Qy 181 LTNNNTGNPKLTQSOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTRQ 232
Db 247 LTNNNTGNPKLTQSOMETFHSSTSEGNLSPSOVSTTSEHPSSPPTTRQ 298

RESULT 5

US-08-467-969A-8
; Sequence 8, Application US/08467969A
; Patent No. 6168786
; GENERAL INFORMATION:
; APPLICANT: Klein, Michel H
; APPLICANT: Du, Run-Pan
; APPLICANT: Ewasyszyn, Mary E
; TITLE OF INVENTION: Chimeric Immunogens
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim & McBurney
; STREET: 330 University Avenue, 6th Floor
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5G 1R7
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,969A
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/001,554
; FILING DATE: 06-JAN-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9200117.1
; FILING DATE: 06-JAN-1992
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REFERENCE/DOCKET NUMBER: 1038-475 MIS:bn
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; TELEX: 065-24567 SIMBAS
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 298 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-467-969A-8

Query Match 100.0%; Score 1229; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 2,4e-105;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 HKVLTFTAIQDATSQIKNTPTPYLTQDPOLGISFSNLSSEITTSQTTTILASTPGVKSML 60

Db 67 HKVLTFAIIODATSOIKNTPTPTVLTODPOLGISFSNLSSEITTSQTTTIIASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCWAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPKPTKPKFTTKKDLKPQTTKPKREVPTTKPEEPTINTTKNTTTT 180
Db 187 KRIPNKKPGKKTTPKPTKPKFTTKKDLKPQTTKPKREVPTTKPEEPTINTTKNTTTT 246
QY 181 LITNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 232
Db 247 LITNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298

RESULT 6
US-08-467-961A-8
; Sequence 8, Application US/08467961A
; Patent No. 6171783

GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasashyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 330 University Avenue, 6TH Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,961A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/001,554
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9200117.1
FILING DATE: 06-JAN-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-476 MIS:dh
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-467-961A-8

Query Match 100.0%; Score 1229; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.4e-105;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HKVLTFAIIODATSOIKNTPTPTVLTODPOLGISFSNLSSEITTSQTTTIIASTTPGVKSNL 60
|||||

Db 67 HKVLTFAIIODATSOIKNTPTPTVLTODPOLGISFSNLSSEITTSQTTTIIASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCWAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPKPTKPKFTTKKDLKPQTTKPKREVPTTKPEEPTINTTKNTTTT 180
Db 187 KRIPNKKPGKKTTPKPTKPKFTTKKDLKPQTTKPKREVPTTKPEEPTINTTKNTTTT 246
QY 181 LITNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 232
Db 247 LITNNTGNPKLTSQMETFHSTSEGNLSPSOVSTTSEHPSQSSPPNTTRQ 298

RESULT 7
US-08-001-554A-8
; Sequence 8, Application US/08001554A
; Patent No. 6225091

GENERAL INFORMATION:
APPLICANT: Klein, Michel H
APPLICANT: Du, Run-Pan
APPLICANT: Ewasashyn, Mary E
TITLE OF INVENTION: Chimeric Immunogens
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Slim & McBurney
STREET: 330 University Avenue, 6th Floor
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5G 1R7
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/001,554A
FILING DATE: 06-JAN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Stewart, Michael I
REGISTRATION NUMBER: 24,973
REFERENCE/DOCKET NUMBER: 1038-286
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 595-1155
TELEFAX: (416) 595-1163
TELEX: 065-24567 SIMBAS
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE:
US-08-001-554A-8

Query Match 100.0%; Score 1229; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 2.4e-105;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HKVLTFAIIODATSOIKNTPTPTVLTODPOLGISFSNLSSEITTSQTTTIIASTTPGVKSNL 60
|||||
Db 67 HKVLTFAIIODATSOIKNTPTPTVLTODPOLGISFSNLSSEITTSQTTTIIASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCWAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKORONKPPKPNNDHFEVFNVPSCISNNPTCWAIC 186
QY 121 KRIPNKKPGKKTTPKPTKPKFTTKKDLKPQTTKPKREVPTTKPEEPTINTTKNTTTT 180
Db 187 KRIPNKKPGKKTTPKPTKPKFTTKKDLKPQTTKPKREVPTTKPEEPTINTTKNTTTT 246

QY 121 KRPNKPKGKTTTKPKPKPKTKKDLKPKQTKPKKEVPTT-----KPTKEPTINT 172
Db 288 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 347
QY 173 TKTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 230
Db 348 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 405

RESULT 14

US-08-928-361B-6
; Sequence 6, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1721 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-361B-6

Query Match 15.3%; Score 187.5; DB 3; Length 1721;
Best Local Similarity 25.6%; Pred. No. 1.1e-08;
Matches 61; Conservative 21; Mismatches 125; Indels 31; Gaps 2;
QY 1 HKVTLTALIDATSOIKNTTPYLYLTOPDQLGISFSNLSELTSTQTTILASTTPGVKSNL 60
Db 191 HTTT 250
QY 61 OPTTKTKNTTTTOTOPOSKPTTKORONKPPNKNNDHFHFVFNVCISCSNPTCAIC 120
Db 251 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 287
QY 121 KRPNKPKGKTTTKPKPKPKTKKDLKPKQTKPKKEVPTT-----KPTKEPTINT 172
Db 288 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 347
QY 173 TKTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 230
Db 348 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 405

Db 348 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 405

RESULT 15

US-08-928-361B-27
; Sequence 27, Application US/08928361B
; Patent No. 6071518
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
; TITLE OF INVENTION: SPECIES INFECTIONS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PETERS, VERNY, JONES & BIRSA
; STREET: 385 Sherman Avenue, Suite 6
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-1840
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentln Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/928,361B
; FILING DATE: 12-SEP-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026,062
; FILING DATE: 13-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: VERNY, Hana
; REGISTRATION NUMBER: 30,518
; REFERENCE/DOCKET NUMBER: 480.76-1(HV)
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-1677
; TELEFAX: 650-324-1678
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 216 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-928-361B-27

Query Match 14.8%; Score 181.5; DB 3; Length 216;
Best Local Similarity 26.0%; Pred. No. 2.6e-09;
Matches 59; Conservative 21; Mismatches 118; Indels 29; Gaps 2;
QY 4 TLTTLAIDATSOIKNTTPYLYLTOPDQLGISFSNLSELTSTQTTILASTTPGVKSNL 63
Db 17 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 76
QY 64 TVTKTKNTTTTOTOPOSKPTTKORONKPPNKNNDHFHFVFNVCISCSNPTCAIC 123
Db 77 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 111
QY 124 PNKPKGKTTTKPKPKPKTKKDLKPKQTKPKKEVPTT-----KPTKEPTINT 183
Db 112 TTKKPTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 167
QY 184 NNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 230
Db 168 TTTTNTTTTLLNNNTGPKRLTSOMETPHSTSSSEGNLSPSOVSTTSEHPSSPPNTT 214

Search completed: May 1, 2003, 13:48:04
Job time: 14.7566 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:46:33 ; Search time 60.4075 Seconds
(without alignments)
331.391 Million cell updates/sec

Title: US-09-462-816-4
Perfect score: 1229
Sequence: 1 HKVTLTIAIQDTSQIKNT.....VSTSEHPSPSPPTTRQ 232

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328255 seqs, 86286685 residues

Total number of hits satisfying chosen parameters: 328255

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*

2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*

3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*

4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*

5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*

6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*

7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*

8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*

9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*

10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*

11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*

12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*

13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*

14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	193	15.7	5179	9	US-10-025-380-1068
2	193	15.7	5179	10	US-09-922-217-1068
3	193	15.7	5179	10	US-09-833-263-1068
4	169.5	13.8	1367	10	US-09-801-368-108
5	162	13.2	941	12	US-10-124-557-14
6	162	13.2	1022	12	US-10-124-557-84
7	162	13.2	1038	12	US-10-124-557-74
8	162	13.2	1049	12	US-10-124-557-58
9	162	13.2	1140	12	US-10-124-557-104
10	162	13.2	1270	12	US-10-124-557-42
11	162	13.2	1311	12	US-10-124-557-42
12	162	13.2	1313	12	US-10-124-557-142
13	162	13.2	1314	12	US-10-124-557-50
14	162	13.2	1320	12	US-10-124-557-46
15	162	13.2	1320	12	US-10-124-557-60
16	162	13.2	1354	12	US-10-124-557-48
17	162	13.2	1361	12	US-10-124-557-40
18	162	13.2	1363	12	US-10-124-557-52
19	162	13.2	1404	12	US-10-124-557-2

20	162	13.2	1404	12	US-10-124-557-62	Sequence 62, App
21	161.5	13.1	288	10	US-09-216-393-341	Sequence 341, App
22	161.5	13.1	288	10	US-09-216-393-344	Sequence 344, App
23	144.5	11.8	2828	9	US-10-176-847-54	Sequence 54, App
24	144.5	11.8	2828	10	US-09-905-129-21	Sequence 21, App
25	144.5	11.8	2828	10	US-09-991-630-21	Sequence 21, App
26	143	11.6	995	9	US-09-984-130-48	Sequence 48, App
27	142.5	11.6	449	9	US-10-028-072-224	Sequence 224, App
28	142.5	11.6	449	9	US-10-121-049-224	Sequence 224, App
29	142.5	11.6	449	9	US-10-123-904-224	Sequence 224, App
30	142.5	11.6	449	9	US-10-140-470-224	Sequence 224, App
31	142.5	11.6	449	9	US-10-175-746-224	Sequence 224, App
32	142.5	11.6	449	9	US-10-176-918-224	Sequence 224, App
33	142.5	11.6	449	9	US-10-177-865-224	Sequence 224, App
34	142.5	11.6	449	9	US-10-140-474-224	Sequence 224, App
35	142.5	11.6	449	9	US-10-142-431-224	Sequence 224, App
36	142.5	11.6	449	9	US-10-143-114-224	Sequence 224, App
37	142.5	11.6	449	9	US-10-140-002-224	Sequence 224, App
38	142.5	11.6	449	9	US-10-142-419-224	Sequence 224, App
39	142.5	11.6	449	9	US-10-123-262-224	Sequence 224, App
40	142.5	11.6	449	9	US-10-142-423-224	Sequence 224, App
41	142.5	11.6	449	9	US-10-121-050-224	Sequence 224, App
42	142.5	11.6	449	9	US-10-141-755-224	Sequence 224, App
43	142.5	11.6	449	9	US-10-143-032-224	Sequence 224, App
44	142.5	11.6	449	9	US-10-123-108-224	Sequence 224, App
45	142.5	11.6	449	9	US-10-123-108-224	Sequence 224, App

ALIGNMENTS

RESULT 1

US-10-025-380-1068

Sequence 1068, Application US/10025380

Publication No. US20020182191A1

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Lodes, Michael J.

APPLICANT: Secrist, Heather

APPLICANT: Benson, Darin R.

APPLICANT: Meagher, Madeleine Joy

APPLICANT: Stolk, John A.

APPLICANT: Wang, Tongtong

APPLICANT: Jiang, Yugu

APPLICANT: Smith, Carole L.

APPLICANT: King, Gordon E.

APPLICANT: Wang, Aijun

APPLICANT: Clapper, Jonathan D.

APPLICANT: Skeiky, Yasir A. W.

APPLICANT: Fanger, Gary R.

APPLICANT: Vedvick Thomas S.

APPLICANT: Carter, Darick

TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS

TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE

FILE REFERENCE: 210121.471C14

CURRENT APPLICATION NUMBER: US/10/025,380

CURRENT FILING DATE: 2001-12-19

NUMBER OF SEQ ID NOS: 1129

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1068

LENGTH: 5179

TYPE: PRT

ORGANISM: Homo sapiens

US-10-025-380-1068

Query Match 15.78; Score 193; DB 9; Length 5179;
Best Local Similarity 30.68; Pred. No. 6.5e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

OY 5 LTFMIAIDA--TSQIKNTTP-----TYIQDPQIGISFNSLSEIT--SQTTLIASTPG 55
DB 1504 MTFPTTPASTITLPPPTTSPPTTTTTPPTTTPSPPTTTPPTTTPPTTTPPTTTP 1563

```
QY 56 ----VKSNIQPTTVTKNTTT-----QTOPSKPTTKORONKPPKNNDHF 98
      :      :      :      :      :      :      :      :      :
Db 1564 PPTTTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1623
      :      :      :      :      :      :      :      :      :
QY 99 FEVENFVPCISCSNNPQCAICKRIPNKKPGKTTTKPKTKPTTKKDLKPQTKPKE 158
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TPPTSTTLPT-----TPSPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
      :      :      :      :      :      :      :      :      :
QY 159 VPTTKTEPTINTKNTNTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTSE 218
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTMTTTPSTTTP--SSPITTTTPSSITTPSPPTTMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 219 HPSQSPSPNTT 230
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 2
US-09-922-217-1068
; Sequence 1068, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-217-1068

Query Match      15.7%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.5e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIIQDA--TSQIKNTTP-----TYLQDPQLGISFSNLSKIT--SQTTLTASTPG 55
      :      :      :      :      :      :      :      :      :
Db 1504 MTPTTPPASTTTLPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1563
      :      :      :      :      :      :      :      :      :
QY 56 ----VKSNIQPTTVTKNTTT-----QTOPSKPTTKORONKPPKNNDHF 98
      :      :      :      :      :      :      :      :      :
Db 1564 PPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1623
      :      :      :      :      :      :      :      :      :
QY 99 FEVENFVPCISCSNNPQCAICKRIPNKKPGKTTTKPKTKPTTKKDLKPQTKPKE 158
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TPPTSTTLPT-----TPSPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
      :      :      :      :      :      :      :      :      :
QY 159 VPTTKTEPTINTKNTNTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTSE 218
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTMTTTPSTTTP--SSPITTTTPSSITTPSPPTTMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 219 HPSQSPSPNTT 230
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 3
```

```
US-09-833-263-1068
; Sequence 1068, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeline J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1068
; LENGTH: 5179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-263-1068

Query Match      15.7%; Score 193; DB 10; Length 5179;
Best Local Similarity 30.6%; Pred. No. 6.5e-06;
Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

QY 5 LTTAIIQDA--TSQIKNTTP-----TYLQDPQLGISFSNLSKIT--SQTTLTASTPG 55
      :      :      :      :      :      :      :      :      :
Db 1504 MTPTTPPASTTTLPTTTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTT 1563
      :      :      :      :      :      :      :      :      :
QY 56 ----VKSNIQPTTVTKNTTT-----QTOPSKPTTKORONKPPKNNDHF 98
      :      :      :      :      :      :      :      :      :
Db 1564 PPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1623
      :      :      :      :      :      :      :      :      :
QY 99 FEVENFVPCISCSNNPQCAICKRIPNKKPGKTTTKPKTKPTTKKDLKPQTKPKE 158
      :      :      :      :      :      :      :      :      :
Db 1624 TPI--TPPTSTTLPT-----TPSPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
      :      :      :      :      :      :      :      :      :
QY 159 VPTTKTEPTINTKNTNTTLLTNNTGNPKLTSOMETPHSTSEGNLSPSOVSTSE 218
      :      :      :      :      :      :      :      :      :
Db 1676 PPTTTPSS--PITTTSPPTTMTTTPSTTTP--SSPITTTTPSSITTPSPPTTMTTP 1731
      :      :      :      :      :      :      :      :      :
QY 219 HPSQSPSPNTT 230
      :      :      :      :      :      :      :      :      :
Db 1732 SPPTTSPPTT 1743
      :      :      :      :      :      :      :      :      :

RESULT 4
US-09-801-368-108
; Sequence 108, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
```

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 1367
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-108

```

```

Query Match          13.8%; Score 169.5; DB 10; Length 1367;
Best Local Similarity 26.2%; Pred. No. 8.5e-05;
Matches 62; Conservative 36; Mismatches 94; Indels 45; Gaps 8;

```

```

QY 18 KNTPTVYLODDPOLGISFSLSEITQTTIIASTTPGV--KSNLOPTVTKKNTT--T 73
DB 209 KSSSTTSSSTSSSTTSSSTSSSTSSSTSSSTSSSTSSSTSSSTSSSTAPATPTTSC 268
QY 74 QTPSKPTTKORONKPPNPNNDHFVEVFNVCISCSNNPCWALCKRIPKPKPKTT 133
DB 269 KEKPTPTTTSCKKEKPTTPHND-----TTPCT-----KKTTTSKTCCTKT 310
QY 134 TKPTKPTTKTKKDLK-----QTKPREVPTTKPTTE-----PTINTTKNTITTL 181
DB 311 TTPVPTSSSTSSAPVPTPSSSTSSAPVTSSTTSSAPVPTPSSSTTSSAPVTS 370
QY 182 LTNNTGNKLTISOMETFTSTSEGNLSPSOVSTTSEHSQP-----SSPNTT 230
DB 371 VTSSTT---ESSAPVTSSTTSSAPVPTPSSSTTSSAPVTSSTTSSAPVTS 424

```

RESULT 5

```

US-10-124-557-14
; Sequence 14, Application US/10124557
; Patent No. US20020137894A1
; GENERAL INFORMATION:

```

```

APPLICANT: Turner, Katherine

```

```

Clark, Stephen C.

```

```

Jacobs, Kenneth

```

```

Hewick, Rodney M.

```

```

Gesner, Thomas G.

```

```

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

```

```

NUMBER OF SEQUENCES: 143

```

```

CORRESPONDENCE ADDRESSES:

```

```

ADDRESS: Genetics Institute, Inc.

```

```

STREET: 87 Cambridgepark Drive

```

```

CITY: Cambridge

```

```

STATE: Massachusetts

```

```

COUNTRY: U.S.A.

```

```

ZIP: 02140

```

```

COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/10/124,557

```

```

FILING DATE: 16-Apr-2002

```

```

CLASSIFICATION: <unknown>

```

```

PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 07/643,502

```

```

FILING DATE: 18-JAN-1991

```

```

APPLICATION NUMBER: US 07/546,114

```

```

FILING DATE: 29-JUN-1990

```

```

APPLICATION NUMBER: US 07/457,196

```

```

FILING DATE: 23-DEC-1989

```

```

APPLICATION NUMBER: US 07/390,901

```

```

FILING DATE: 08-AUG-1989

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: Cserr, Luann

```

```

REGISTRATION NUMBER: 31,822

```

```

REFERENCE/DOCKET NUMBER: GT 5190

```

```

TELECOMMUNICATION INFORMATION:

```

```

TELEPHONE: (617)876-1170

```

```

TELEFAX: (617)876-5851

```

```

INFORMATION FOR SEQ ID NO: 14:

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 941 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-10-124-557-14

```

```

Query Match          13.2%; Score 162; DB 12; Length 941;
Best Local Similarity 24.6%; Pred. No. 0.00021;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

```

```

QY 6 TTAIIODATSOIKNTPTVYLODDPOLGISFSLSEITQTTIIASTTPGVKSNLOPTTV 65
DB 125 TSKVLAKPTPKAEITTK-----GPAITPKPEPTTTPKEASTTP---KEPTPTT 172
QY 66 K-----TKNTTQTOTOPSKPTTKORONKPPNPNNDHFVEVFNVCISCSN 112
DB 173 KSAPTPKPEAPVTTTKSAPTTPKEAPVTTTKPEAPVTTTKPEP-----APTTPKEP 221
QY 113 NPTCAIACKRIPN-----KKPGKTTTKP-----TKKPTFKTK 146
DB 222 APTTTKSAPTPKPEAPVTTTKSAPTTPKEAPVTTTKPEAPVTTTKPEAPVTTTK 281
QY 147 KDL-----KPTTKPREVPTTKP-----TEEPINTTKNTITTTLLTNNTGNPKLTS 194
DB 282 EPAPTPAKKPAPTTPKPEAPVTTTKPEAPVTTTKPEAPVTTTKPEAPVTTTKPEAP 341
QY 195 OMETFTSTSSGNLSPSOVSTTSEHSOPSPPNTT 230
DB 342 TTKSAPTTPKE-----PSPPTTKPEAPVTTTKPEAPVTT 373

```

RESULT 6

```

US-10-124-557-84
; Sequence 84, Application US/10124557
; Patent No. US20020137894A1
; GENERAL INFORMATION:

```

```

APPLICANT: Turner, Katherine

```

```

Clark, Stephen C.

```

```

Jacobs, Kenneth

```

```

Hewick, Rodney M.

```

```

Gesner, Thomas G.

```

```

TITLE OF INVENTION: Megakaryocyte Stimulating Factors

```

```

NUMBER OF SEQUENCES: 143

```

```

CORRESPONDENCE ADDRESSES:

```

```

ADDRESS: Genetics Institute, Inc.

```

```

STREET: 87 Cambridgepark Drive

```

```

CITY: Cambridge

```

```

STATE: Massachusetts

```

```

COUNTRY: U.S.A.

```

```

ZIP: 02140

```

```

COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

CURRENT APPLICATION DATA:

```

```

APPLICATION NUMBER: US/10/124,557

```

```

FILING DATE: 16-Apr-2002

```

```

CLASSIFICATION: <unknown>

```

```

PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 07/643,502

```

```

FILING DATE: 18-JAN-1991

```

```

APPLICATION NUMBER: US 07/546,114

```

```

FILING DATE: 29-JUN-1990

```

```

APPLICATION NUMBER: US 07/457,196

```

```

FILING DATE: 23-DEC-1989

```

```

APPLICATION NUMBER: US 07/390,901

```

```

FILING DATE: 08-AUG-1989

```

```

ATTORNEY/AGENT INFORMATION:

```

```

NAME: Cserr, Luann

```

```

REGISTRATION NUMBER: 31,822

```


CURRENT APPLICATION DATA:

ZIP: 02140
COMPUTER READABLE FORM:


```

      TITLE OF INVENTION: Megakaryocyte Stimulating Factors
      NUMBER OF SEQUENCES: 143
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Genetics Institute, Inc.
      STREET: 87 Cambridgepark Drive
      CITY: Cambridge
      STATE: Massachusetts
      COUNTRY: U.S.A.
      ZIP: 02140
      COMPUTER READABLE FORM:
      MEDIUM TYPE: floppy disk
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/10/124,557
      FILING DATE: 16-Apr-2002
      CLASSIFICATION: <Unknown>
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/643,502
      FILING DATE: 18-JAN-1991
      APPLICATION NUMBER: US 07/546,114
      FILING DATE: 29-JUN-1990
      APPLICATION NUMBER: US 07/457,196
      FILING DATE: 29-DEC-1989
      APPLICATION NUMBER: US 07/390,901
      FILING DATE: 08-AUG-1989
      ATTORNEY/AGENT INFORMATION:
      NAME: Cserr, Luann
      REGISTRATION NUMBER: 31,822
      REFERENCE/DOCKET NUMBER: GI 5190
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617)876-1170
      TELEFAX: (617)876-5851
      INFORMATION FOR SEQ ID NO: 142:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1313 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      SEQUENCE DESCRIPTION: SEQ ID NO: 142:
US-10-124-557-142

Query Match      13.2%; Score 162; DB 12; Length 1313;
Best Local Similarity 24.6%; Pred. No. 0.0003;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

      6 TTAIIQDATSQIKNTPTTYLTDDPOLGISFSNLSEITSGTTIIASTPGVKSNIQPTTV 65
      233 TSKVLAKPTPKAEITTK-----GPAITPKPEPTPTTKREPASTP---KEPTPTTI 280
      66 K-----TKNTTTOTOPSKPTTKQROKNPKPNNDHFVFNFPSCISN 112
      281 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKREPAPTTKPEP-----APTTPKEP 329
      113 NPTCAIICKRIPN-----KKPGKKTTPK-----TKKPTFKTK 146
      330 APTTTKSAPTTPKEPAPTTKREPAPTTKREPAPTTKREPAPTTKREPAPTTK 389
      147 KDL-----KPQTKPKREVPTTKP-----TEEPTINTTKNTNTTLLTNNTTGNPKLTS 194
      390 EPAPTAPKPAPTTPKREPAPTTKREPAPTTKREPAPTTKREPAPTTKREPAPTT 449
      195 QMETFHSISSEGNLSPSOVSTTSEHSOPSSPNTT 230
      450 TTKSAPTTPKPKE-----PSPPTTKREPAPTTKREPAPTT 481

```

RESULT 13

US-10-124-557-50
 Sequence 50, Application US/1014557
 Patent No. US20020137894A1
 GENERAL INFORMATION:

```

      APPLICANT: Turner, Katherine
      Clark, Stephen C.
      Jacobs, Kenneth
      Hewick, Rodney M.
      Gesner, Thomas G.
      TITLE OF INVENTION: Megakaryocyte Stimulating Factors
      NUMBER OF SEQUENCES: 143
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Genetics Institute, Inc.
      STREET: 87 Cambridgepark Drive
      CITY: Cambridge
      STATE: Massachusetts
      COUNTRY: U.S.A.
      ZIP: 02140
      COMPUTER READABLE FORM:
      MEDIUM TYPE: floppy disk
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/10/124,557
      FILING DATE: 16-Apr-2002
      CLASSIFICATION: <Unknown>
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: US 07/643,502
      FILING DATE: 18-JAN-1991
      APPLICATION NUMBER: US 07/546,114
      FILING DATE: 29-JUN-1990
      APPLICATION NUMBER: US 07/457,196
      FILING DATE: 29-DEC-1989
      APPLICATION NUMBER: US 07/390,901
      FILING DATE: 08-AUG-1989
      ATTORNEY/AGENT INFORMATION:
      NAME: Cserr, Luann
      REGISTRATION NUMBER: 31,822
      REFERENCE/DOCKET NUMBER: GI 5190
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617)876-1170
      TELEFAX: (617)876-5851
      INFORMATION FOR SEQ ID NO: 50:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 1314 amino acids
      TYPE: amino acid
      TOPOLOGY: linear
      MOLECULE TYPE: protein
      SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-10-124-557-50

Query Match      13.2%; Score 162; DB 12; Length 1314;
Best Local Similarity 24.6%; Pred. No. 0.0003;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;

      6 TTAIIQDATSQIKNTPTTYLTDDPOLGISFSNLSEITSGTTIIASTPGVKSNIQPTTV 65
      234 TSKVLAPTPKAEITTK-----GPAITPKPEPTPTTKREPASTP---KEPTPTTI 281
      66 K-----TKNTTTOTOPSKPTTKQROKNPKPNNDHFVFNFPSCISN 112
      282 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKREPAPTTKPEP-----APTTPKEP 330
      113 NPTCAIICKRIPN-----KKPGKKTTPK-----TKKPTFKTK 146
      331 APTTTKSAPTTPKEPAPTTKREPAPTTKREPAPTTKREPAPTTKREPAPTTK 390
      147 KDL-----KPQTKPKREVPTTKP-----TEEPTINTTKNTNTTLLTNNTTGNPKLTS 194
      391 EPAPTAPKPAPTTPKREPAPTTKREPAPTTKREPAPTTKREPAPTTKREPAPTT 450
      195 QMETFHSISSEGNLSPSOVSTTSEHSOPSSPNTT 230
      451 TTKSAPTTPKPKE-----PSPPTTKREPAPTTKREPAPTT 482

```

RESULT 14
US-10-124-557-46
; Sequence 46, Application US/10124557
; Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 46:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-10-124-557-46
Query Match 13.2%; Score 162; DB 12; Length 1320;
Best Local Similarity 24.6%; Pred. No. 0.0003;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;
QY 6 TTAIIQDATSQIKNTTPYTLTODPOLGISFSNLSITSTTTTILASTTPGVKSNLOPTTV 65
DB 240 TSKVLAKPPTKPAETTK-----GPAITPKKEPTPTTPKPAETTP--KEPTPTTI 287
QY 66 K-----TKNNTTQTOPSKPTTKORONKPNKPNNDPFHEVENFVPCISCSN 112
DB 288 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTPKEP-----APTTTKEP 336
QY 113 NPTCWAICKRIPN-----KKPGKTTTKP-----TKKPTTKTK 146
DB 337 APPTTKSAPTTPKEPAPTTKSAPTTPKEPAPTTPKEPAPTTPKEPAPTTPK 396
QY 147 KDL-----KQPTTKPKKEVPTTKP-----TEEPTINTTKNTITTTLLNNTTGNFKLTS 194
DB 397 EPAPTAPKRPAPPTTPKEPAPTTKKEPAPTTKPSPTTPKEPAPTTKSAPTTKKEPAPT 456

QY 195 QMETFHTSSEGNLSPSOVSTSEHPSOPSPSPNTT 230
DB 457 TTKSAPTTPKE-----PSPPTTKKEPAPTTPKKEPAPTT 488
RESULT 15
US-10-124-557-60
; Sequence 60, Application US/10124557
; Patent No. US20020137894A1
GENERAL INFORMATION:
APPLICANT: Turner, Katherine
Clark, Stephen C.
Jacobs, Kenneth
Hewick, Rodney M.
Gesner, Thomas G.
TITLE OF INVENTION: Megakaryocyte Stimulating Factors
NUMBER OF SEQUENCES: 143
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/124,557
FILING DATE: 16-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/643,502
FILING DATE: 18-JAN-1991
APPLICATION NUMBER: US 07/546,114
FILING DATE: 29-JUN-1990
APPLICATION NUMBER: US 07/457,196
FILING DATE: 29-DEC-1989
APPLICATION NUMBER: US 07/390,901
FILING DATE: 08-AUG-1989
ATTORNEY/AGENT INFORMATION:
NAME: Geert, Luann
REGISTRATION NUMBER: 31,822
REFERENCE/DOCKET NUMBER: GI 5190
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)876-1170
TELEFAX: (617)876-5851
INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 1320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 60:
US-10-124-557-60
Query Match 13.2%; Score 162; DB 12; Length 1320;
Best Local Similarity 24.6%; Pred. No. 0.0003;
Matches 68; Conservative 26; Mismatches 104; Indels 78; Gaps 9;
QY 6 TTAIIQDATSQIKNTTPYTLTODPOLGISFSNLSITSTTTTILASTTPGVKSNLOPTTV 65
DB 240 TSKVLAKPPTKPAETTK-----GPAITPKKEPTPTTPKPAETTP--KEPTPTTI 287
QY 66 K-----TKNNTTQTOPSKPTTKORONKPNKPNNDPFHEVENFVPCISCSN 112
DB 288 KSAPTTPKEPAPTTKSAPTTPKEPAPTTKKEPAPTTPKEP-----APTTTKEP 336
QY 113 NPTCWAICKRIPN-----KKPGKTTTKP-----TKKPTTKTK 146
DB 337 APPTTKSAPTTPKEPAPTTKSAPTTPKEPAPTTPKEPAPTTPKEPAPTTPK 396

OY 147 KDL-----KQQTTPKKEVPPTKP-----TEEPTINTTKNTITTTLLNNNTGNDPKLTS 194
Db 397 EPAPTPAPKPPAPPTTPKKEPAPPTTPKKEPAPTTTKEPSPTTPKKEPAPTTTKSAPTTTKEPAPPT 456
OY 195 QMETFHSTSSKGNLSPSOVSTTSEHPSQSPSPNTT 230
Db 457 TTKSAPTTPKE---PSPTTKEPAPPTTPKKEPAPTT 488

Search completed: May 1, 2003, 13:56:40
Job time : 66.4075 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:41:33 ; Search time 15.3208 Seconds
(Without alignments)
1455.750 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKYTLTTAIIODATSOIKNT.....VSTTSEHPSPSSPNTTRQ 232

Scoring table:

BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1216	98.9	298	1	MGNZRL
2	1120	91.1	298	1	MGNZRL
3	1052	85.6	297	2	JO1208
4	1045	85.0	297	2	JO1205
5	1042	84.8	298	2	JO1206
6	1037	84.4	297	2	JO1207
7	1033	84.1	297	2	JO1209
8	1019	82.9	298	2	JO1207
9	1009	82.1	297	2	JO1204
10	505.5	41.1	292	1	MGNZ60
11	497	40.4	292	1	MGNZ18
12	425	34.6	467	1	VHNZ
13	202	16.4	307	1	GSFP3
14	193	15.7	3020	2	A43932
15	191	15.5	662	2	A45155
16	190	15.5	332	2	A53715
17	188	15.3	1832	2	T31113
18	186.5	15.2	851	2	T22696
19	185.5	15.1	379	2	S50125
20	183	14.9	217	2	S01358
21	183	14.9	1161	2	S57180
22	179	14.6	354	2	T46740
23	178	14.5	250	2	MGNZBR
24	177	14.4	257	1	MGNZBR
25	175	14.2	770	2	T22808
26	175	14.2	825	2	T29634
27	172.5	14.0	337	2	S20074
28	171.5	14.0	216	2	I51920
29	170.5	13.9	263	2	JO2284

30	170.5	13.9	279	2	S53363	mucin 5AC (clone J
31	169.5	13.8	1367	1	S44478	glican 1.4-alpha-g
32	164.5	13.4	2476	2	T34022	zonadhesin - pig
33	163.5	13.3	294	2	A37232	mucin, tracheal (A
34	163.5	13.3	402	2	E86185	hypothetical prote
35	163	13.3	371	2	S20075	promastigote surfa
36	163	13.3	477	2	S53362	mucin 5AC (clone J
37	160.5	13.1	393	2	S63335	771-7 protein - fr
38	160	13.0	263	2	S01360	salivary glue prot
39	159.5	13.0	263	2	JO2388	glycoprotein G - o
40	158	12.9	1777	2	T34369	hypothetical prote
41	157.5	12.8	1630	2	A53577	ascites staloglyco
42	157	12.8	248	2	P00769	glycoprotein G - b
43	156.5	12.7	263	2	A48732	attachment glycop
44	155	12.6	1118	2	A48292	mucin, tracheobron
45	153.5	12.5	3507	2	T34513	hypothetical prote

ALIGNMENTS

RESULT 1

Major surface glycoprotein G - human respiratory syncytial virus (strain Long)

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999

C:Accession: A32703; S12279

R:Johnson, P. R.; Spriggs, M. K.; Olmsted, R. A.; Collins, P. L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

A:Reference number: A32703; MUID:87289657; PMID:2441388

A:Accession: A32703

A:Molecule type: mRNA

A:Residues: 1-298 <JOH>

A:Cross-references: GB:M17212; NID:9333940; PIDN:AAA47411.1; PID:9333941

R:Garcia-Barrero, B.; Portela, A.; Delgado, T.; Lopez, J.A.; Melero, J.A.

EMBO J. 9, 4181-4187, 1990

A:Title: Frame shift mutations as a novel mechanism for the generation of neutralizat

A:Reference number: S12279; MUID:91065351; PMID:2249671

A:Accession: S12279

A:Molecule type: mRNA

A:Residues: 1-298 <GAR>

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted <TMN>

F:65,103,135,179,237,250,251,273,294/Binding site: carbohydrate (Asn) (covalent) #sta

Query Match

Best local Similarity 98.9%; Score 1216; DB 1; Length 298;

Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	HKYTLTTAIIODATSOIKNTPTTYLTQDPOLGTSFNTSEISQTTILASTTPGVKSL	60
DB	67	HKYTLTTAIIODATSOIKNTPTTYLTQDPOLGTSFNTSEISQTTILASTTPGVKSL	126
QY	61	QPTTVTKNTTQTOPSKPTTKORONKPPNPFHEVEFNFCSSNNPTCAIC	120
DB	127	QPTTVTKNTTQTOPSKPTTKORONKPPNPFHEVEFNFCSSNNPTCAIC	186
QY	121	KRIPNKPGKTTTKPTKKPTKTKKDKLPQTKRKEVPTTKPEEPITNTTKINITTT	180
DB	187	KRIPNKPGKTTTKPTKKPTKTKKDKLPQTKRKEVPTTKPEEPITNTTKINITTT	246
QY	181	LTNNNTGNPKLTISOMETHSTSEGNLSPOVSTTSEHPSPSSPNTTRQ	232
DB	247	LTNNNTGNPKLTISOMETHSTSEGNLSPOVSTTSEHPSPSSPNTTRQ	298

RESULT 2

Major surface glycoprotein G - human respiratory syncytial virus

C:Species: human respiratory syncytial virus

C:Date: 28-May-1986 #sequence_revision 28-May-1986 #text_change 24-Sep-1999

RESULT 9

J01204 attachment protein - human respiratory syncytial virus (strain RSB642)

N:Alternate names: G protein

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Feb-1997

C:Accession: J01204

R:Canv, P.A.; Matthews, D.A.; Pringle, C.R.

J:Gene, Virol. 72, 2091-2096, 1991

A:Title: Identification of variable domains of the attachment (G) protein of subgroup A

A:Reference number: J01204; MUID:91374005; PMID:1895054

A:Accession: J01204

A:Molecule type: mRNA

A:Residues: 1-297 <CAN>

C:Comment: Respiratory syncytial virus commonly causes severe lower respiratory tract ill

children and adults.

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:85,103,135,144,237,273/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 82.1%; Score 1009; DB 2; Length 297;

Matches 194; Conservative 11; Mismatches 26; Indels 0; Gaps 0;

QY 1 HKVLTITAIQDATSQIKNTPTTYLQDPQGISFSNLSSEITTSQTTTILASTPGVKSML 60

DB 67 HRTVSTTTIIQDATNQIKNTPTTYLQNPGLGISPNPSBITLTITIDPTTPGVKLT 126

QY 61 QPTTVTKNTTTTQTOPSKPTTKQKRONKPPKNNDFHEFVFNVCISCSNNPTCMAIC 120

DB 127 QSTTVTKNTTTTQTOPSKPTTKQKRONKPPKNNDFHEFVFNVCISCSNNPTCMAIC 186

QY 121 KRIPNKKPKGKTTTKPTTKPTTKKDLKPOTKPREVPTTPEPTINTKTNTTT 180

DB 187 KRIPNKKPKGKTTTKPTTKPTTKKDLKPOTKPREVPTTPEPTINTKTNTTT 246

QY 181 LTLNNTGNPKLTSQMETFHSSTSEGNLSPSQVSTSEHPSQSPSPNTR 231

DB 247 PLTNSNTRNPELTSQMETFHSSTSEGNLSPSQVSTSEHPSQSPSPNTR 297

RESULT 10

MGWZ60

major surface glycoprotein G - human respiratory syncytial virus (strain 8/60)

N:Alternate names: attachment glycoprotein G

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999

C:Accession: A37077

R:Sullender, W.M.; Anderson, K.; Wertz, G.W.

Virology 178, 195-203, 1990

A:Title: The respiratory syncytial virus subgroup B attachment glycoprotein: analysis of

group subgroup virus challenge.

A:Reference number: A37077; MUID:90357765; PMID:1697126

A:Accession: A37077

A:Molecule type: mRNA

A:Residues: 1-292 <SUL>

A:Cross-references: EMBL:M55633; NID:g333944; PIDN:AAA47413.1; PID:g333945

C:Genetics:

A:Gene: G

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:45-63/Domain: transmembrane #status predicted <TMN>

F:81,86,100,230,290/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 41.1%; Score 505.5; DB 1; Length 292;

Matches 108; Conservative 28; Mismatches 88; Indels 3; Gaps 3;

QY 1 HKVLTITAIQDATSQIKNTPTTYLQDPQGISFSNLSSEITTSQTTTILASTPGVKSML 60

DB 67 HKVLTITVYQITKNHTGKNISTYLTQVPERVNSKOPTTSPHTNSATISPNKSET 126

QY 61 QPTTVTKNTTTTQTOPSKPTTKQKRONKPPKNNDFHEFVFNVCISCSNNPTCMAIC 120
DB 127 HHTTAQTKGRITTSQTKPKPTTKSKRSKPKKPDYHFEVFNVCISGNMOLCKSIC 186
QY 121 KRIPNKKPKGKTTTKPTTKPTTKKDLKPOTKPREVPTTPEPTINTKTNTTT 179
DB 187 KRIPNKKPKGKTTTKPTTKPTTKKDLKPOTKPREVPTTPEPTINTKTNTTT 246
QY 180 TLLNNTGNPKLTSQMETFHSSTSEGNLSPSQVSTSEHPSQSPSPNTR 226
DB 247 SQSTVLDITTPKTYLQDQSLHSTSENPSSSTQIPITASE-PS-TSNP 251

RESULT 11

MGWZ18

major surface glycoprotein G - human respiratory syncytial virus (strain 18537)

C:Species: human respiratory syncytial virus

C:Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 16-Jul-1999

C:Accession: B32703

R:Johnson, P.R.; Spriggs, M.K.; Olmsted, R.A.; Collins, P.L.

Proc. Natl. Acad. Sci. U.S.A. 84, 5625-5629, 1987

A:Title: The G glycoprotein of human respiratory syncytial viruses of subgroups A and

F:41-63/Domain: transmembrane #status predicted <TMN>

F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted

A:Accession: B32703

A:Molecule type: mRNA

A:Residues: 1-292 <JOH>

A:Cross-references: GB:M1213; NID:g333942; PIDN:AAA47412.1; PID:g333943

C:Superfamily: respiratory syncytial virus major surface glycoprotein G

C:Keywords: glycoprotein; transmembrane protein

F:41-63/Domain: transmembrane #status predicted <TMN>

F:81,86,100/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 40.4%; Score 497; DB 1; Length 292;

Matches 104; Conservative 28; Mismatches 88; Indels 2; Gaps 2;

QY 1 HKVLTITAIQDATSQIKNTPTTYLQDPQGISFSNLSSEITTSQTTTILASTPGVKSML 60

DB 67 HKVLTITVYQITKNHTGKNISTYLTQVPERVNSKOPTTSPHTNSATISPNKSET 126

QY 61 QPTTVTKNTTTTQTOPSKPTTKQKRONKPPKNNDFHEFVFNVCISCSNNPTCMAIC 120

DB 127 HHTTAQTKGRITTSQTKPKPTTKSKRSKPKKPDYHFEVFNVCISGNMOLCKSIC 186

QY 121 KRIPNKKPKGKTTTKPTTKPTTKKDLKPOTKPREVPTTPEPTINTKTNTTT 179

DB 187 KRIPNKKPKGKTTTKPTTKPTTKKDLKPOTKPREVPTTPEPTINTKTNTTT 246

QY 180 TLLNNTGNPKLTSQMETFHSSTSEGNLSPSQVSTSEHPS 221

DB 247 SQSTVLDITTPKTYLQDQSLHSTSENPSSSTQIPITASE-PS 287

RESULT 12

VHNZ

nucleocapsid protein (version 2) - human respiratory syncytial virus

C:Species: human respiratory syncytial virus

C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999

C:Accession: A04026

R:Elango, N.; Venkatesan, S.

Nucleic Acids Res. 11, 5941-5951, 1983

A:Title: Amino acid sequence of respiratory syncytial virus capsid protein.

A:Reference number: A04026; MUID:83299261; PMID:6310521

A:Accession: A04026

A:Molecule type: mRNA

A:Residues: 1-467 <ELA>

A:Cross-references: GB:X00001; NID:g61215; PIDN:CAA24906.1; PID:g61216

C:Genetics:

A:Gene: N

C:Superfamily: respiratory syncytial virus nucleocapsid protein

C:Keywords: nucleocapsid

Query Match

34.6%; Score 425; DB 1; Length 467;

Db 1564 PPPTTTTTPPTTSPPTTTTSPPTTTTTPPTTTTTPPTTTTTPPTTTTTPPTTTTSPPTT 1623
QY 99 FEVNFVPCISICNNPTCWAICRIIPNKKPGKTTTKPKPKTKKDKLPQTKPKE 158
Db 1624 TPI--TPTSTTTLPPT-----TTPSPPTTTTTPPTTTTTPSPPTTTTTPPTTTTTP 1675
QY 159 VPTTKPTEPINTTKITTTTLLNNTGNPKLISQMEFHSTSSSEGNLSPSOVSTSE 218
Db 1676 PPTTPSS--PITTPSPPTTMTTPSPPTTP--SSPITTTTPSSSTTTTSPPTTMTTP 1731
QY 219 HPSQSSPPTT 230
Db 1732 SPPTTSPPTT 1743

RESULT 15

A45155
mucin FIM-C.1 - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 26-May-1994 #sequence_revision 26-May-1994 #text_change 21-Jul-2000
C:Accession: A45155
R:Hauser, F.; Hoffmann, W.
J. Biol. Chem. 267, 24620-24624, 1992
A:Title: P-domains as shuffled cysteine-rich modules in integumentary mucin C.1 (FIM-C.1)
A:Reference number: A45155; MUID:93077556; PMID:1447205
A:Accession: A45155
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-662 <HAU>
A:Cross-References: GB:L02115; NID:9214147; PIDN:AAA74725.1; PID:9951460
C:Superfamily: trefoil homology
F:162-202/Domain: trefoil homology <TRP1>
F:307-347/Domain: trefoil homology <TRP2>
F:354-394/Domain: trefoil homology <TRP3>
F:526-566/Domain: trefoil homology <TRP4>
F:573-613/Domain: trefoil homology <TRP5>
F:621-661/Domain: trefoil homology <TRP6>

Query Match

15.5%; Score 191; DB 2; Length 662;

Best Local Similarity 29.0%; Pred. No. 1, 2e-05;
Matches 80; Conservative 19; Mismatches 111; Indels 66; Gaps 12;

QY 2 KYLTATAIQDATSOIKNTT-PTYLQDPQIGISFNSLSEITSQTTTIIASTTGVKSNL 60
Db 240 KATPTTTTATKATPTTTTATKATTTTP-----TATTTTATKATTTPTTTTTP 287
QY 61 QPTTVATKNTTTQT-----QPSKP-----TTKORNK-----PNK----- 92
Db 288 TPPTTTTATKATTTTTSGECKMPSKREDGYSGITESQCRKGGCCFDSISIPQTKMCFYT 347
QY 93 --PNDPFHEVENFVPCSI-----CSNNPTC-----WAICKRIIPNKKPGKTTT 134
Db 348 LSGVACCKEYPSQVDCGRGITADCCROKNCQCFDSISIGTKW--CFYSTQVAATKTTT 405
QY 135 KPTKDPTEFKTKDKLPQTKPKREVPPTKPTBEPINTTKINTTTTLLNNTGNPKLIS 194
Db 406 TPPTTTT-PTTTTATKATTTTPTTTTPTTTTPTTTTPTTTTATKATTTPTTTPTTTKAT 464
QY 195 QMETFHSTSSSEGNLSPSOVSTSEHPSQSSPPTT 230
Db 465 TPPT--TPTTPTTTTATKATTT--PTTTTPTTTT 496

Search completed: May 1, 2003, 13:47:17
Job time : 16.3208 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 12:35:37 : Search time 10.0679 Seconds

(Without alignments)
955.755 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVLTALITADATSOIKNT.....VSTSEHPSPSPNTTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1216	98.9	298	1	VGIG_HRSVL
2	1120	91.1	298	1	VGIG_HRSVA
3	1052	85.6	297	1	VGIG_HRSV6
4	1045	85.0	297	1	VGIG_HRSV3
5	1037	84.4	297	1	VGIG_HRSV7
6	1033	84.1	297	1	VGIG_HRSV4
7	1019	82.9	298	1	VGIG_HRSV5
8	1009	82.1	297	1	VGIG_HRSV2
9	505.5	41.1	292	1	VGIG_HRSV8
10	497	40.4	292	1	VGIG_HRSV1
11	202	16.4	307	1	SGS3_DROME
12	198.5	16.2	2700	1	ZAN_HUMAN
13	193	16.7	5179	1	MOC2_HUMAN
14	191	15.5	662	1	MOC1_XENLA
15	183	14.9	217	1	SGS3_DROSI
16	183	14.9	1161	1	DAN4_YEAST
17	180.5	14.7	263	1	DAN4_YEAST
18	179.5	14.6	263	1	VGIG_HRSV1
19	177	14.4	257	1	VGIG_HRSV4
20	176.5	14.4	263	1	VGIG_HRSV3
21	170.5	13.9	263	1	VGIG_HRSV2
22	169.5	13.8	1367	1	AMYH_YEAST
23	169	13.8	257	1	VGIG_YEAST
24	166	13.5	2282	1	ZAN_RABIT
25	164.5	13.4	2476	1	ZAN_PIG
26	160	13.0	263	1	SGS3_DROYA
27	159.5	13.0	263	1	VGIG_ORSVW
28	159	12.9	257	1	VGIG_BRSV5
29	156	12.7	257	1	VGIG_BRSV4
30	154	12.5	257	1	VGIG_BRSV3
31	150	12.2	3178	1	KS89_CAEEL
32	147.5	12.0	797	1	VGIX_HSEVB
33	147.5	12.0	907	1	VGPI3_EBV

34	147	12.0	860	1	CH12_COICM
35	145.5	11.8	5376	1	ZAN_MOUSE
36	144.5	11.8	1260	1	ALSI_CANAL
37	143	11.6	886	1	VGPI3_EBV8
38	141	11.5	1251	1	YOUN_CAEEL
39	139.5	11.4	400	1	MUOL_XENLA
40	139	11.3	767	1	AMYH_SACDI
41	138	11.2	457	1	P60_LISIN
42	137.5	11.2	681	1	VGPI3_MABVM
43	135.5	11.0	768	1	AMYH_SACDI
44	135	11.0	605	1	YHCB_YEAST
45	134.5	10.9	596	1	YAAC_SCHPO

ALIGNMENTS

```

RESULT 1
VGIG_HRSVL
ID VGIG_HRSVL STANDARD: PRT: 298 AA.
AC P20895;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (subgroup A / strain Long).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11260;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87289657; PubMed=2441388;
RA Johnson P.R., Spriggs M.K., Olmsted R.A., Collins P.L.;
RT "The G glycoprotein of human respiratory syncytial viruses of
RT subgroups A and B: extensive sequence divergence between
RT antigenically related proteins."
RL Proc. Natl. Acad. Sci. U.S.A. 84:5625-5629(1987).
CC -!- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -!- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -!- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to: license@sib-sib.ch).
CC -----
DR EMBL: M17212; AAA47411.1; -.
DR PIR: A32703; MGN2RL.
DR InterPro: IPR000925; Glycoprot G.
DR Pfam: PF00802; Glycoprotein G: 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37
FT TRANSMEM 38 66
FT DOMAIN 67 298
FT CARBOHYD 103 103 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 250 250 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 294 294 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 298 AA; 32781 MW; B79EFA4BA73B0E CRC64;

Query Match 98.9%; Score 1216; DB 1; Length 298;
Best Local Similarity 99.1%; Pred. No. 1.3e-75;
Matches 230; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 HKVLTALITADATSOIKNTPTLYLADPOLGISFNSLSEITSOITTLASTTPGVKSNL 60
CC      |
CC      |
CC      |
Db 67 HKVLTALITADATSOIKNTPTLYLADPOLGISFNSLSEITSOITTLASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 120
CC      |
CC      |
Db 127 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 186
QY 121 KRIPIKKPKGKTTTKPTKPKPTKTKDLPQTTPKPKVPTTKPEPTINTKTNITTT 180
CC      |
CC      |
Db 187 KRIPIKKPKGKTTTKPTKPKPTKTKDLPQTTPKPKVPTTKPEPTINTKTNITTT 246
QY 181 LLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 232
Db 247 LLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 2
VGLG_HRSVA STANDARD; PRT; 298 AA.
ID VGLG_HRSVA
AC P03423;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain A2).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11259;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85216636; PubMed=3858865;
RA Wertz G.W., Collins P.L., Huang Y., Gruber C., Levine S., Ball L.A.;
RT "Nucleotide sequence of the G protein gene of human respiratory
RT syncytial virus reveals an unusual type of viral membrane protein.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4075-4079(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=86067198; PubMed=4069997;
RA Satake M., Colligan J.E., Elango N., Norrby E., Venkatesan S.;
RT "Respiratory syncytial virus envelope glycoprotein (G) has a novel
RT structure.";
RL Nucleic Acids Res. 13:7795-7812(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=95266253; PubMed=7747420;
RA Connors M., Crowe J.E. Jr., Firestone C.Y., Murphy B.R., Collins P.L.;
RT "A cold-passaged, attenuated strain of human respiratory syncytial
RT virus contains mutations in the F and L genes.";
RL Virology 208:478-484(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97187925; PubMed=9035372;
RA Crowe J.E. Jr., Firestone C.Y., Whitehead S.S., Collins P.L.,
RA Murphy B.R.;
RT "Acquisition of the ts phenotype by a chemically mutagenized cold-
RT passaged human respiratory syncytial virus vaccine candidate results
RT from the acquisition of a single mutation in the polymerase (L)
RT gene.";
RL Virus Genes 13:269-273(1996).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M1486; AAB59857.1; -
DR EMBL: X03149; CAA26928.1; -
DR EMBL: U50362; AAB86653.1; -
DR EMBL: U50363; AAB86675.1; -
DR EMBL: U63644; AAC55969.1; -
DR PIR: A04039; MGNZ.
DR InterPro: IPR000925; Glycoprot-G.
DR Pfam: PF00802; Glycoprotein-G; 1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 66
FT DOMAIN 67 298 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 251 251 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 298 AA; 32586 MW; 993C3D2DD68BC634 CRC64;

Query Match 91.1%; Score 1120; DB 1; Length 298;
Best Local Similarity 92.7%; Pred. No. 4e-69; 12; Indels 0; Gaps 0;
Matches 215; Conservative 5; Mismatches 12;

QY 1 HKVLTALITADATSOIKNTPTLYLADPOLGISFNSLSEITSOITTLASTTPGVKSNL 60
Db 67 HKVLTALITADATSOIKNTPTLYLADPOLGISFNSLSEITSOITTLASTTPGVKSNL 126
QY 61 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 120
Db 127 OPTTVKTKNTTTTQTOPSKPTTKOROKPKPKNNDDHFEVFNVPSCISNNPTCAIC 186
QY 121 KRIPIKKPKGKTTTKPTKPKPTKTKDLPQTTPKPKVPTTKPEPTINTKTNITTT 180
Db 187 KRIPIKKPKGKTTTKPTKPKPTKTKDLPQTTPKPKVPTTKPEPTINTKTNITTT 246
QY 181 LLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 232
Db 247 LLTNNTGNPKLTSOMETFHSTSEGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 3
VGLG_HRSV6 STANDARD; PRT; 297 AA.
ID VGLG_HRSV6
AC P27025;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsh6256).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11256;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT "Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.";
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
CC -----
DR PIR: J01208; J01208.
DR InterPro: IPR000925; Glycoprot-G.
DR Pfam: PF00802; Glycoprotein-G; 1.

```

KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37
FT TRANSMEM 38 66
FT DOMAIN 67 297
FT CARBOHYD 103 103
FT CARBOHYD 135 135
FT CARBOHYD 237 237
FT CARBOHYD 251 251
FT CARBOHYD 294 294
SQ SEQUENCE 297 AA: 32708 MW: 6781756C38B64A80 CRC64;

Query Match
Best Local Similarity 85.6%; Score 1052; DB 1; Length 297;
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVLTALITQDATSQIKNTPTTYLTQDPLGIFSNLSEITTSQTTILASTTPGVKSNL 60
DB 67 HKVLTALITQDATSQIKNTPTTYLTQDPLGIFSNLSEITTSQTTTPPTPSAESTP 126
QY 61 OPTTVKTKNTTQTOPSKPTTKORONKPPNNDFHEVFENFVPCISCSNPTCWAIC 120
DB 127 QSTTVKTKNTTQTOPSKPTTKORONKPPNNDFHEVFENFVPCISCSNPTCWAIC 186
QY 121 KRIPNKKPGKTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNTTT 180
DB 187 KRIPNKKPGKTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNTTT 246
QY 181 LTTNNTGNPKLTISQMETFHSTSSSEGNLSPSOVSTTSEHPSPSSPPTT 230
DB 247 LTTNNTGNPKLTISQMETFHSTSSSEGNLSPSOVSTTSEHPSPSSPPTT 296

RESULT 4
VGIG_HRSV3
ID VGIG_HRSV3 STANDARD; PRT; 297 AA.
AC P27022;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb1734).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11253;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT *Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.*;
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR: J01205; J01205.
DR InterPro: IPR000925; Glycoprot-G.
DR Pfam: PF00802; Glycoprotein_G_1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37
FT TRANSMEM 38 66
FT DOMAIN 67 297
FT CARBOHYD 135 135
FT CARBOHYD 237 237
FT CARBOHYD 251 251
SQ SEQUENCE 297 AA: 32525 MW: 48448F9E091E1802 CRC64;

Query Match
Best Local Similarity 85.0%; Score 1045; DB 1; Length 297;
Matches 199; Conservative 6; Pred. No. 4.5e-64;

Matches 200; Conservative 10; Mismatches 21; Indels 0; Gaps 0;

QY 1 HKVLTALITQDATSQIKNTPTTYLTQDPLGIFSNLSEITTSQTTILASTTPGVKSNL 60
DB 67 HKVLTALITQDATSQIKNTPTTYLTQDPLGIFSNLSEITTSQTTTPPTPSAESTP 126
QY 61 OPTTVKTKNTTQTOPSKPTTKORONKPPNNDFHEVFENFVPCISCSNPTCWAIC 120
DB 127 QSTTVKTKNTTQTOPSKPTTKORONKPPNNDFHEVFENFVPCISCSNPTCWAIC 186
QY 121 KRIPNKKPGKTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNTTT 180
DB 187 KRIPNKKPGKTTKPTKPKPTTKKDLKPQTTPKREVPPTKPEEPTINTKTNTTT 246
QY 181 LTTNNTGNPKLTISQMETFHSTSSSEGNLSPSOVSTTSEHPSPSSPPTT 231
DB 247 LTTNNTGNPKLTISQMETFHSTSSSEGNLSPSOVSTTSEHPSPSSPPTT 297

RESULT 5
VGIG_HRSV7
ID VGIG_HRSV7 STANDARD; PRT; 297 AA.
AC P27026;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-AUG-1992 (Rel. 23, Last annotation update)
DE Major surface glycoprotein G (Attachment glycoprotein G).
GN G.
OS Human respiratory syncytial virus (strain rsb6614).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11257;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91374005; PubMed=1895054;
RA Cane P.A., Matthews D.A., Pringle C.R.;
RT *Identification of variable domains of the attachment (G) protein of
RT subgroup A respiratory syncytial viruses.*;
RL J. Gen. Virol. 72:2091-2096(1991).
CC -1- FUNCTION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE
CC RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND
CC HEMAGGLUTININATING ACTIVITIES.
CC -1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED
CC CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.
CC -1- PTM: MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE CHAINS
CC DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.
DR PIR: J01209; J01209.
DR InterPro: IPR000925; Glycoprot-G.
DR Pfam: PF00802; Glycoprotein_G_1.
KW Transmembrane; Glycoprotein.
FT DOMAIN 1 37
FT TRANSMEM 38 66
FT DOMAIN 67 297
FT CARBOHYD 103 103
FT CARBOHYD 135 135
FT CARBOHYD 237 237
FT CARBOHYD 251 251
FT CARBOHYD 294 294
SQ SEQUENCE 297 AA: 32670 MW: 58B384028E437ACD CRC64;

Query Match
Best Local Similarity 84.4%; Score 1037; DB 1; Length 297;
Matches 199; Conservative 5; Pred. No. 1.6e-63;

DR PIR: A03329; GSFF3.
 DR Flybase: FBgn0003373; Sgs3.
 KW Repeat: signal.
 FT SIGNAL 1 23 POTENTIAL
 FT CHAIN 24 307 SALIVARY GLUE PROTEIN SCS-3.
 SQ SEQUENCE 307 AA; 32196 MW; 45803DEDD16C418BC CRC64;
 Query Match 16.4%; Score 202; DB 1; Length 307;
 Best Local Similarity 32.8%; Pred. No. 3.4e-07;
 Matches 77; Conservative 17; Mismatches 93; Indels 48; Gaps 13;

QY 4 TLTATITQATISQIKNTTPPYLQDDPOLGISFNSLSEITISQTTTILASTTPGV--KSNLQ 61
 DB 49 TTTTTCAPPTQO--STTQPPCTTSKP-----TTKQTQTQTPCTTPPTTKATTK 97
 QY 62 PTTVATKNTTTTQOPSKPTTKORONKPPKPNNDHFEVFNVPVCSGNNPTCAICK 121
 DB 98 PTTTKATTTKATTTKPT--TTKQTTQLP-----CTTPPTTKQTTT 136
 QY 122 RINKKP--GKKTTPK--TKKPTFKTK--KDLKQTTKREVPPTK--TEEPINTTKN 176
 DB 137 QLEPCTPTTKPTTKPTTKPTKPT--TTKPTTKPTTKP---TTTKPTTKPT--TTKPT 189
 QY 177 ITTTLTNNNTGKPKLTSQMEFHTSSSEGNLSPSOVSTSEHPSPSSPPTTR 231
 DB 190 TTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPTTKPT--TTKPTTK 242

RESULT 12
 ZAN_HUMAN STANDARD; PRT: 2700 AA.
 ID ZAN_HUMAN
 AC Q9Y493; 000218;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE zonadhesin (Fragment).
 GN ZAN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE OF 1-2379 FROM N.A.
 RX MEDLINE=99018118; PubMed=9799793;
 RA Glocker G., Scherer S., Schatzevov R., Boright A., Weber J.,
 RA Tsui L.C., Rosenchal A.;
 RT "Large-scale sequencing of two regions in human chromosome 7q22:
 RT analysis of 650 kb of genomic sequence around the EPO and CUTL1 loci
 RT reveals 17 genes.";
 RT Genome Res. 8:1060-1073(1998).
 RN [2]
 RP SEQUENCE OF 2338-2700 FROM N.A.
 RC TISSUE=Testis;
 RA MEDLINE=97271566; PubMed=9126492;
 RA Gao Z., Hartum T., Garbers D.L.;
 RT "Chromosome localization of the mouse zonadhesin gene and the human
 RT zonadhesin gene (ZAN).";
 RT Genomics 41:119-122(1997).
 RL Genomics 41:119-122(1997).
 CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR
 CC SIGNALING.
 CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE
 CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).
 CC -1- TISSUE SPECIFICITY: IN TESTIS; PRIMARILY IN HAPLOID SPERMATIDS.
 CC -1- DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE
 CC ZONA PELLUCIDA.
 CC -1- DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS,
 CC THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF
 CC SPERMATOZOA OR PROMOTING ADHESION TO THE OVIDUCTAL ISTHMUS.
 CC -1- DOMAIN: THE WPD DOMAINS 2 AND 3 MAY MEDIATE COVALENT
 CC OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUC2).
 CC -1- SIMILARITY: CONTAINS 3.5 MAM DOMAINS.

CC -1- SIMILARITY: CONTAINS 4.5 WPD DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF053356; AAC78790.1;
 CC DR EMBL: U81191; AAC51208.1;
 CC DR Genew: HGNC:12857; ZAN.
 CC DR MIM: 602372;
 CC DR InterPro: IPR000561; EGF-like.
 CC DR InterPro: IPR000998; MAM domain.
 CC DR InterPro: IPR002919; TIL Cysrich.
 CC DR InterPro: IPR003328; TIL Cysrich.
 CC DR InterPro: IPR001846; WPD_D.
 CC DR Pfam: PF00094; WPD; 4.
 CC DR Pfam: PF00629; MAM; 4.
 CC DR Pfam: PF01826; TIL; 5.
 CC DR Pfam: PF02345; TILa; 4.
 CC DR SMART: SM00216; WPD; 1.
 CC DR PROSITE: PS01186; EGF_2; 3.
 CC DR PROSITE: PS00740; MAM_1; 1.
 CC DR PROSITE: PS00600; MAM_2; 4.
 CC DR Glycoprotein; Transmembrane;
 CC E1 NON_TER 1
 CC E1 DOMAIN <1 109 MAM 1.
 CC E1 DOMAIN 112 136 MAM 2 (PARTIAL).
 CC E1 DOMAIN 161 326 MAM 3.
 CC E1 DOMAIN 322 446 MAM 4.
 CC E1 DOMAIN 483 951 66 X HEPTAPEPTIDE REPEATS (APPROXIMATE)
 CC E1 DOMAIN 953 1065 (MUCIN-LIKE DOMAIN).
 CC E1 DOMAIN 1066 1454 WPD 2.
 CC E1 DOMAIN 1455 1861 WPD 3.
 CC E1 DOMAIN 1862 2292 WPD 4.
 CC E1 DOMAIN 2293 2684 WPD 5.
 CC E1 DOMAIN ? ? EGF-LIKE.
 CC E1 CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 403 403 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 1023 1023 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 1099 1099 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 1618 1618 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 1737 1737 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 1832 1832 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 1878 1878 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 2136 2136 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CARBOHYD 2505 2505 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 CONFLICT 2374 2379 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC E1 NON_TER 2700 NNOKMA -> RAGGCP (IN REF. 1).
 CC SO SEQUENCE 2700 AA; 293013 MW; 80E60CC0B12277B1 CRC64;
 Query Match 16.2%; Score 198.5; DB 1; Length 2700;
 Best Local Similarity 29.5%; Pred. No. 5.3e-06;
 Matches 74; Conservative 34; Mismatches 96; Indels 47; Gaps 12;

QY 17 IKNTPTTYLQDDPOLGISFNSLSEITISQTTTILASTTPGVKSNLQPTTVKTKNTTQTO 76
 DB 531 IPSKPTTILTEKPTIP--SEKPTIPSEKPTI--STEKPTVPTEPPT--PTEETTYMEB 584
 QY 77 PSKPTTK-----ORONKPPKPNNDHFEVFNVPVCSGNNPTCAICKRIPNK----- 127
 DB 585 PVLPTEKPSIPTKPSIPTKPSIPTKPSISMEETLISIEKPTICPEKPTIPTEKPTIPTEKSTIS 644
 QY 128 GKKTIT-----TKPTKPKPTFKTKDL--KPTTKPK--VPTTK--PTEPTINTTK 174
 DB 645 PEKPTPTTEKPTIPTEKPTISTEKPTIPTEK--TISPEKLTIPTEKLTIPTEKPTIPTEK 703
 QY 175 TLTNTT-----TLTNNNTGKPKLTSQMEFHTSSSEGNLSPSOVST-----T 216


```
FT CARBOHYD 4616 4616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4627 4627 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4752 4752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4787 4787 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4881 4881 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4888 4888 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4955 4955 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 4970 4970 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5019 5019 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5038 5038 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 5069 5069 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 1351 1351 H -> L (IN REF. 3).
FT CONFLICT 1412 1412 T -> S (IN REF. 3).
FT CONFLICT 1449 1449 L -> P (IN REF. 3).
FT CONFLICT 1504 1504 M -> T (IN REF. 3).
FT CONFLICT 4192 4192 G -> S (IN REF. 2).
SQ SEQUENCE 5179 AA; 540295 MM; 85CD57JFB9A5663 CRC64;
```

Query Match 15.78; Score 193; DB 1; Length 5179;
Best Local Similarity 30.68; Pred. No. 2.4e-05;

Matches 77; Conservative 23; Mismatches 114; Indels 38; Gaps 9;

```
QY 5 LTTAIIQDA--TSQIKNTTP-----TYLNDDPOLGISFSLSEIT--SOTTTTLASTTGG 55
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1504 MTRPIIPASTTLPPTTSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1563
QY 56 ----VKSNIQPTTVTKTKNTTT-----OTOPSKPTTKOKRNKPKPNKPNNDPH 98
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1564 PPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1623
QY 99 FEVENVPQSGSNNTCAICRIKPKKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK 158
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1624 TPT--TPTTSTTTLPTT-----TTPSPPTTTPPTTTPPTTTPPTTTPPTTTPPTTTP 1675
QY 159 VPTTKPTTEPTINTTKNTTLLTNNTGNPKTSOMEPHSTSEGMSPSQVSTTSE 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1676 PPTTTPSS--PITTPPSPTTTPPTTTPPTTTP--SSPITTTTPSSSTTTPSPPTTTPPTTTP 1731
QY 219 HPSQSPSPNTT 230
   : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 1732 SPPTTTPSPPTTTP 1743
```

```
RESULT 14
MUC1_XENLA STANDARD: PRT; 662 AA.
AC 005049;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integumentary mucin C.1 (FIM-C.1) (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8335;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3; 4; 5; 6 AND 7).
RC TISSUE=Skin;
RX MEDLINE=93077556; PubMed=1447205;
RA Hauser F., Hoffmann W.;
RT "p-domains as shuffled cysteine-rich modules in integumentary mucin
RT C.1 (FIM-C.1) from Xenopus laevis. Polydispersity and genetic
RT polymorphism.";
RL J. Biol. Chem. 267:24620-24624(1992).
CC -1- FUNCTION: COULD BE INVOLVED IN DEFENSE AGAINST MICROBIAL
CC INFECTIONS. PROTECTS THE EPITHELIA FROM EXTERNAL ENVIRONMENT.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- ALTERNATIVE PRODUCTS: At least 7 isoforms: 1 (shown here), 2, 3,
CC 4, 5, 6 and 7; may be produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: SKIN.
CC -1- PTM: EXTENSIVELY O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 6 P-TYPE (TREPOLL) DOMAINS.
```

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L02115; AAA74725.1; -.
CC PIR: A45155; A45155.
CC DR HSP: P01359; 2PSP.
CC DR InterPro: IPR000519; P_trefoll.
CC DR Pfam: PF00088; trefoll; 6.
CC DR PROSITE: PS00025; P_TREPOLL; 6.
CC DR Repeat: Amphibian skin; Glycoprotein; Alternative splicing.
CC KW NON_TER
CC FT 1 144
CC FT 81 144 8 X 8 AA APPROXIMATE TANDEM REPEATS,
CC FT DOMAIN Ala/Thr-Rich.
CC FT REPEAT 1-1.
CC FT REPEAT 81 88 1-2.
CC FT REPEAT 89 96 1-3.
CC FT REPEAT 97 104 1-4.
CC FT REPEAT 105 112 1-5.
CC FT REPEAT 113 120 1-6.
CC FT REPEAT 121 128 1-7.
CC FT REPEAT 129 136 1-8.
CC FT REPEAT 137 144 1-9.
CC FT REPEAT 161 202 8 X APPROXIMATE TANDEM REPEATS, THR-RICH.
CC FT DOMAIN 1.
CC FT REPEAT 218 224 2-1.
CC FT REPEAT 225 239 2-2.
CC FT REPEAT 240 249 2-3.
CC FT REPEAT 250 259 2-4.
CC FT REPEAT 260 275 2-5.
CC FT REPEAT 276 287 2-6.
CC FT REPEAT 288 294 2-7.
CC FT REPEAT 295 301 2-8.
CC FT REPEAT 306 347 P-TYPE 2.
CC FT DOMAIN 3.
CC FT 353 394 P-TYPE 3.
CC FT 402 522 12 X APPROXIMATE TANDEM REPEATS,
CC FT THR-RICH.
CC FT REPEAT 3-1.
CC FT REPEAT 412 411 3-2.
CC FT REPEAT 420 431 3-3.
CC FT REPEAT 432 443 3-4.
CC FT REPEAT 444 453 3-5.
CC FT REPEAT 454 460 3-6.
CC FT REPEAT 461 472 3-7.
CC FT REPEAT 473 479 3-8.
CC FT REPEAT 480 491 3-9.
CC FT REPEAT 492 498 3-10.
CC FT REPEAT 499 515 3-11.
CC FT REPEAT 516 522 3-12.
CC FT DOMAIN 4.
CC FT 525 566 P-TYPE 4.
CC FT DOMAIN 5.
CC FT 572 613 P-TYPE 5.
CC FT DOMAIN 6.
CC FT 620 661 P-TYPE 6.
CC FT DISULFID 162 188 BY SIMILARITY.
CC FT DISULFID 172 187 BY SIMILARITY.
CC FT DISULFID 182 199 BY SIMILARITY.
CC FT DISULFID 307 333 BY SIMILARITY.
CC FT DISULFID 317 332 BY SIMILARITY.
CC FT DISULFID 327 344 BY SIMILARITY.
CC FT DISULFID 354 380 BY SIMILARITY.
CC FT DISULFID 364 379 BY SIMILARITY.
CC FT DISULFID 374 391 BY SIMILARITY.
CC FT DISULFID 526 532 BY SIMILARITY.
CC FT DISULFID 536 551 BY SIMILARITY.
CC FT DISULFID 546 563 BY SIMILARITY.
CC FT DISULFID 573 599 BY SIMILARITY.
CC FT DISULFID 583 598 BY SIMILARITY.
CC FT DISULFID 593 610 BY SIMILARITY.
CC FT DISULFID 621 647 BY SIMILARITY.
```

FT DISULFID 631 646 BY SIMILARITY.
FT DISULFID 641 658 BY SIMILARITY.
FT VASAPLIC 240 259 MISSING (IN ISOFORMS 5 AND 6).
FT VASAPLIC 250 259 MISSING (IN ISOFORMS 4 AND 7).
FT VASAPLIC 276 294 MISSING (IN ISOFORMS 4 AND 6).
FT VASAPLIC 278 278 MISSING (IN ISOFORM 5).
FT VASAPLIC 306 350 MISSING (IN ISOFORMS 2 AND 5).
FT VASAPLIC 420 498 MISSING (IN ISOFORM 3).
FT VARIANT 276 276 K -> E.
FT VARIANT 354 354 K -> E.
FT VARIANT 415 415 C -> R.
SQ SEQUENCE 662 AA; 67774 MW; F08527F1ED2FD40 CRC64;

Query Match 15.5%; Score 191; DB 1; Length 662;
Best Local Similarity 29.0%; Pred. No. 4.1e-06;
Matches 80; Conservative 19; Mismatches 111; Indels 66; Gaps 12;

OY 2 KVTLTALIQDTSIKNTT-PTVLTODPOLGISPSNLSITISOTTTILASTTPGVKSNL 60
DB 240 KATPTTTTATKATPTTTTATTTTTP-----TTTTTTKATTTPTTTT 287
OY 61 OPTTVKTKNTTTTQT-----QPSKP-----TKORONK-----PNK----- 92
DB 288 TPTTTTATTTTSTTSECKMEPSKREDGYSGITESCCKGCEPDSIPOTKWCFT 347
OY 93 --PNNDFHEFVFNFPVPCSI-----CSNNPFC-----MAICKRIIPNKKPKGKTTT 134
DB 348 ISQVADCKVEPSQRYDCRGFTTADOCROKNCCEPDSISGTRK--CYSTISQVAAATKTTT 405
OY 135 KPTKPKPTTKKDKLPQTTKREVPPTKPTKEPTINTKTNTITTTLLJNNTGKPKLTS 194
DB 406 TPTTTTATTTTPTTTTATTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTTPTTTT 464
OY 195 QMETFHSTISSEGNLSPSOVSTSEHPQSPPPTT 230
DB 465 TPTT--TTTTPTTTTATTTT--PTTTTPTTTT 496

RESULT 15
SGS3_DROSI STANDARD; PRT; 217 AA.
ID SGS3_DROSI
AC P13729;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Salivary glue protein Sgs-3 precursor.
GN SGS3.
OS Drosophila simulans (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID-7240;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-8833296; PubMed-3138416;
RA Martin C.H., Mayeda C.A., Meyerowitz E.M.;
RT "Evolution and expression of the Sgs-3 glue gene of Drosophila.";
RT J. Mol. Biol. 201:273-287(1988).
CC -!- DEVELOPMENTAL STAGE: PRODUCED BY THIRD-INSTAR LARVAE.
DR PIR, S01358; S01358
DR Flybase, FBgn0012853; Dsim\SGS3.
KW Repeat; Signal.
FT SIGNAL 1 23
FT CHAIN 24 217 SALIVARY GLUE PROTEIN SGS-3.
SQ SEQUENCE 217 AA; 22750 MW; D29894E34025781 CRC64;

Query Match 14.9%; Score 183; DB 1; Length 217;
Best Local Similarity 33.3%; Pred. No. 4.6e-06;
Matches 55; Conservative 10; Mismatches 68; Indels 32; Gaps 5;

OY 33 ISFSNLSSEI-----TSQFTTILASTTPGVKSNLQPTTVKTKNTTTTQT---QPSKPTTK 83
DB 15 IGFANVANCSDCGCPTKATTTTCAPTKPKCKSTTTTTTTTTTTTTTTTTTTTTRAPPTKPTCK 74

OY 84 QRO-----NKPNNKNNDFHEFVFNFPVPCISCSNNPTCMAICKRIIPNKKGKTTTKPT 137
DB 75 STSTTTTTRAPPTKPT-----CKSTSTTTTTRAPPTTCKTSTTTTTT 119
OY 138 KPTFKPTTKKDKLPQTTKREVPPT--TKPTKEPTINTKTNTITTT 180
DB 120 HKPTTHTPTKPTKPTKHTTEPKTKPTKHTTEPKTKPTKHTTPTTTTTTT 164

Search completed: May 1, 2003, 13:45:04
Job time : 12.0679 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 1, 2003, 13:27:43 : Search time 28.0151 Seconds
(without alignments)
1706.328 Million cell updates/sec

Title: US-09-462-816-4

Perfect score: 1229

Sequence: 1 HKVTLTALIQDATSIKNT.....VSTTSEHPSSPPPTTRQ 232

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: SP archaea:*
2: SP bacteria:*
3: SP fungi:*
4: SP human:*
5: SP invertebrate:*
6: SP mammal:*
7: SP mhc:*
8: SP organelle:*
9: SP phage:*
10: SP plant:*
11: SP rodent:*
12: SP virus:*
13: SP vertebrate:*
14: SP unclassified:*
15: SP virus:*
16: SP bacteriap:*
17: SP archaep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1229	100.0	298	12	Q82068	082068 human respi
2	1229	100.0	298	12	Q01929	001929 human respi
3	1093	88.9	298	12	Q09719	009719 respiratory
4	1093	88.9	298	12	Q09634	009634 human respi
5	1088	88.5	279	12	Q09678	099678 human respi
6	1072	87.2	279	12	Q09600	099600 human respi
7	1067	86.8	297	12	Q09VB5	099VB5 human respi
8	1066	86.7	297	12	Q08206	082066 human respi
9	1063	86.5	293	12	Q09VB3	099VB3 human respi
10	1062	86.4	298	12	Q09VB4	099VB4 human respi
11	1056	85.9	297	12	Q08207	082071 human respi
12	1055	85.8	278	12	Q09673	099673 human respi
13	1053	85.7	292	12	Q09VB2	099VB2 human respi
14	1053	85.7	297	12	Q08208	082058 human respi
15	1052	85.6	295	12	Q08359	086359 respiratory
16	1047	85.2	297	12	Q82067	Q82067 human respi

17	1047	85.2	297	12	Q91947	Q91947 human respi
18	1045	85.0	278	12	Q06071	Q06071 human respi
19	1045	85.0	295	12	Q06360	Q06360 respiratory
20	1044	84.9	279	12	Q09680	Q09680 human respi
21	1044	84.9	279	12	Q09689	Q09689 human respi
22	1044	84.9	293	12	Q09VB1	Q09VB1 human respi
23	1043	84.9	298	12	Q08205	Q08205 human respi
24	1041	84.7	278	12	Q09601	Q09601 human respi
25	1040	84.6	297	12	Q82074	Q82074 human respi
26	1040	84.6	297	12	Q82064	Q82064 human respi
27	1039	84.5	279	12	Q09652	Q09652 human respi
28	1039	84.5	292	12	Q09VB0	Q09VB0 human respi
29	1037	84.4	279	12	Q096S9	Q096S9 human respi
30	1037	84.4	295	12	Q82063	Q82063 human respi
31	1037	84.4	297	12	Q08206	Q08206 human respi
32	1036	84.3	278	12	Q09674	Q09674 human respi
33	1036	84.3	279	12	Q096S6	Q096S6 human respi
34	1035	84.2	279	12	Q09679	Q09679 human respi
35	1035	84.2	299	12	Q08207	Q08207 human respi
36	1033	84.1	279	12	Q096S7	Q096S7 human respi
37	1033	84.1	295	12	Q08357	Q08357 respiratory
38	1032	84.0	279	12	Q096S1	Q096S1 human respi
39	1032	84.0	297	12	Q09194	Q09194 human respi
40	1032	84.0	297	12	Q09VC8	Q09VC8 human respi
41	1031	83.9	298	12	Q08206	Q08206 human respi
42	1029	83.7	297	12	Q08207	Q08207 human respi
43	1028	83.6	297	12	Q82079	Q82079 human respi
44	1028	83.6	297	12	Q82057	Q82057 human respi
45	1028	83.6	298	12	Q82062	Q82062 human respi

ALIGNMENTS

RESULT 1

ID Q82068 PRELIMINARY; PRT; 298 AA.
AC Q82068;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Glycoprotein (Fragment).
GN G.
OS Human respiratory syncytial virus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxID=11250;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SUBGROUP A;
RX MEDLINE=9435057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbiza J., Fabrisile S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno E.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein".
RL J. Virol. 68:5448-5459(1994).
DR EMBL: Z33429; CA83872.1; -
DR InterPro: IPR000925; Glycoprot_G.
DR InterPro: IPR003860; Ppantact-act.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
FT NON_TER
SO SQUENCE 298 AA; 32753 MW; 70D80307897A772B CRC64;

Query Match 100.0%; Score 1229; DB 12; Length 298;
Best Local Similarity 100.0%; Pred. No. 6.5e-85;
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HKVTLTALIQDATSIKNTPTYLTPDQPLGSPNSLSEITSTTTILASTTPGVKSNL 60
DB 67 HKVTLTALIQDATSIKNTPTYLTPDQPLGSPNSLSEITSTTTILASTTPGVKSNL 126

QY	61	OPTVTKRNTTQTOPSPKPTTKORONKPKPNNDHFEVFNVPSCISNNPTCAIC	120		
Db	127	OPTVTKRNTTQTOPSPKPTTKORONKPKPNNDHFEVFNVPSCISNNPTCAIC	186		
QY	121	KRIINRKGKTKTKPTKPKPTTKKDKJQTKRKPKEVPTTKPTEBPTINTTKNTT	180		
Db	187	KRIINRKGKTKTKTKPKPKPTTKKDKJQTKRKPKEVPTTKPTEBPTINTTKNTT	246		
OY	181	LLTNNTGPNKPLTSOMETFHSTSEGNLSPSOVSTTSHPSQSPSPPTTQ	232		
Db	247	LLTNNTGPNKPLTSOMETFHSTSEGNLSPSOVSTTSHPSQSPSPPTTQ	298		
RESULT 2					
ID	001929	PRELIMINARY;	PRT; 298 AA.		
AC	001929;				
DT	01-NOV-1996 (Tremblrel. 01, Created)				
DT	01-NOV-1996 (Tremblrel. 01, Last sequence update)				
DT	01-MAR-2002 (Tremblrel. 20, Last annotation update)				
DE	Major surface glycoprotein G (Attachment glycoprotein G).				
GN	G.				
OS	Human respiratory syncytial virus (subgroup A / strain Long).				
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;				
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.				
OX	NCBI_TaxID=11260;				
RN	[1]				
RX	SEQUENCE FROM N.A. PubMed-2249671.				
RA	Medline-91065351; Garcia-Barreno B., Portela A., Delgado T., Lopez J.A., Melero J.A.;				
RT	"Frame shift mutations as a novel mechanism for the generation of				
RT	neutralization resistant mutants of human respiratory syncytial				
RT	virus.";				
RL	EMBO J. 9:4181-4187(1990).				
CC	-1- EDUCATION: UNLIKE THE OTHER PARAMYXOVIRUS ATTACHMENT PROTEINS, THE				
CC	RESPIRATORY SYNCYTIAL VIRUS G PROTEIN LACKS BOTH NEURAMINIDASE AND				
CC	HEMAGGLUTININATING ACTIVITIES.				
CC	-1- SUBCELLULAR LOCATION: EXPRESSED ON THE SURFACE OF THE INFECTED				
CC	CELLS AND INCORPORATED IN THE MEMBRANE OF THE VIRIONS.				
CC	-1- PTM: THIS PROTEIN MAY CARRY 40-80 SEPARATE O-LINKED CARBOHYDRATE				
CC	CHAINS DISTRIBUTED AMONG THE 91 SERINE AND THREONINE RESIDUES.				
DR	EMBL; X17085; CA334937.1; -				
DR	InterPro: IPR000925; Glycoprote_G.				
DR	InterPro: IPR003880; Pantne_atlch.				
DR	Pfam: PF00802; Glycoprotein G; 1.				
DR	PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.				
KM	Transmembrane; Glycoprotein.				
FT	DOMAIN 1 37				
FT	TRANSMEM 38 66				
FT	DOMAIN 67 298				
FT	DOMAIN 205 211				
FT	SITE 64 65				
FT	SITE 74 75				
FT	CARBOHYD 85 85				
FT	CARBOHYD 103 103				
FT	CARBOHYD 250 250				
FT	CARBOHYD 251 251				
FT	CARBOHYD 273 273				
FT	CARBOHYD 284 294				
SO	SEQUENCE 298 AA; 32771 MW; F8CD4213D97C2952 CRC64;				
Query Match 100.0%; Score 1229; DB 12; Length 298;					
Best Local Similarity 100.0%; Pred. No. 6.5e-85;					
Matches 232; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
OY	1	HKVLTITAIIDATSOIKNTPTLYLTODPOLGISFSNLSITSSQTTIIASTTPGVKSNL	60		
Db	67	HKVLTITAIIDATSOIKNTPTLYLTODPOLGISFSNLSITSSQTTIIASTTPGVKSNL	126		
OY	61	OPTVTKRNTTQTOPSPKPTTKORONKPKPNNDHFEVFNVPSCISNNPTCAIC	120		
Db	127	OPTVTKRNTTQTOPSPKPTTKORONKPKPNNDHFEVFNVPSCISNNPTCAIC	186		

DE Attachment protein (g).
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 CC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=52;
 RX MEDLINE=97181512; PubMed=9032893;
 RA Tolley K.P., Marriott A.C., Simpson A., Ploos D.J., Matthews D.A.,
 R Longhurst S.J., Evans J.E., Johnson J.L., Cane P.A., Easton A.J.,
 RA Pringle G.R.;
 RT "Identification of mutations contributing to the reduced virulence of
 RT a modified strain of respiratory syncytial virus.";
 RL Vaccine 14:1637-1646(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=52;
 RA Easton A.J.;
 RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
 DR EMBL: U39662; AAC57026.1; -;
 DR InterPro: IPR000925; Glycoprot.G.
 DR InterPro: IPR003880; Pnantine_attach.
 DR Pfam: PF00802; Glycoprotein_G.1.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
 SQ SEQUENCE 298 AA; 32749 MW; 60F27B29D497F31 CRC64;

Query Match 88.9%; Score 1093; DB 12; Length 298;
 Best Local Similarity 90.1%; Pred. No. 11e-74;
 Matches 209; Conservative 6; Mismatches 17; Indels 0; Gaps 0;

QY 1 HKVLTITAIIDATSOIKNTPTLYLTDPOLGISFSNLSSETTSQTTTILASTPGVKNL 60
 DB 67 HKVLTITAIIDATSOIKNTPTLYLTDPOLGISFSNLSSETTSQTTTILASTPGVKNL 126
 QY 61 OPTTVKNTNTTQOTOPSKPTTKORONKPPKPNNDHFEEVFNFPVCSICSNPTGMAIC 120
 DB 127 QSTTVTKNTNTTQOTOPSKPTTKORONKPPKPNNDHFEEVFNFPVCSICSNPTGMAIC 186
 QY 121 KRIPNKKPGKKTTPKPKPTTKKDKLPQTTKPKKEVPTTKPEEPTINTTKNTTT 180
 DB 187 KRIPNKKPGKKTTPKPKPTTKKDKLPQTTKPKKEVPTTKPEEPTINTTKNTTT 246
 QY 181 LTNNTTGNPKLTSGMETFHSTSGGNLSPSOVSTTSEHPSPSPPTTRQ 232
 DB 247 LTNNTTGNPKLTSGMETFHSTSGGNLSPSOVSTTSEHPSPSPPTTRQ 298

RESULT 5
 Q906T8 PRELIMINARY; PRT; 279 AA.
 AC Q906T8;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein (Fragment).
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 CC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=92011;
 RX MEDLINE=20283719; PubMed=10823752;
 RA Choi E.H., Lee H.J.;
 RT "Genetic diversity and molecular epidemiology of the G protein of
 RT subgroups A and B of respiratory syncytial virus isolated over 9
 RT consecutive epidemics in Korea.";
 RL J. Infect. Dis. 181:1547-1556(2000).
 DR EMBL: AF193308; AAF23731.1; -;
 DR InterPro: IPR000925; Glycoprot.G.

DR InterPro: IPR003880; Pnantine_attach.
 DR Pfam: PF00802; Glycoprotein_G.1.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 279 AA; 30498 MW; 4E1CD2FE7E569E3A CRC64;

Query Match 88.5%; Score 1088; DB 12; Length 273;
 Best Local Similarity 90.1%; Pred. No. 2.3e-74;
 Matches 209; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTITAIIDATSOIKNTPTLYLTDPOLGISFSNLSSETTSQTTTILASTPGVKNL 60
 DB 48 HKVLTITAIIDATSOIKNTPTLYLTDPOLGISFSNLSSETTSQTTTILASTPGVKNL 107
 QY 61 OPTTVKNTNTTQOTOPSKPTTKORONKPPKPNNDHFEEVFNFPVCSICSNPTGMAIC 120
 DB 108 LPTTVKNTNTTQOTOPSKPTTKORONKPPKPNNDHFEEVFNFPVCSICSNPTGMAIC 167
 QY 121 KRIPNKKPGKKTTPKPKPTTKKDKLPQTTKPKKEVPTTKPEEPTINTTKNTTT 180
 DB 168 KRIPNKKPGKKTTPKPKPTTKKDKLPQTTKPKKEVPTTKPEEPTINTTKNTTT 227
 QY 181 LTNNTTGNPKLTSGMETFHSTSGGNLSPSOVSTTSEHPSPSPPTTRQ 232
 DB 228 LTNNTTGNPKLTSGMETFHSTSGGNLSPSOVSTTSEHPSPSPPTTRQ 279

RESULT 6
 Q906U0 PRELIMINARY; PRT; 279 AA.
 AC Q906U0;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Glycoprotein (Fragment).
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 CC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=91242;
 RX MEDLINE=20283719; PubMed=10823752;
 RA Choi E.H., Lee H.J.;
 RT "Genetic diversity and molecular epidemiology of the G protein of
 RT subgroups A and B of respiratory syncytial virus isolated over 9
 RT consecutive epidemics in Korea.";
 RL J. Infect. Dis. 181:1547-1556(2000).
 DR EMBL: AF193306; AAF23729.1; -;
 DR InterPro: IPR000925; Glycoprot.G.
 DR InterPro: IPR003880; Pnantine_attach.
 DR Pfam: PF00802; Glycoprotein_G.1.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 279 AA; 30544 MW; 6B270AEB1CB3533 CRC64;

Query Match 87.2%; Score 1072; DB 12; Length 273;
 Best Local Similarity 90.0%; Pred. No. 3.7e-73;
 Matches 207; Conservative 4; Mismatches 19; Indels 0; Gaps 0;

QY 1 HKVLTITAIIDATSOIKNTPTLYLTDPOLGISFSNLSSETTSQTTTILASTPGVKNL 60
 DB 48 HKVLTITAIIDATSOIKNTPTLYLTDPOLGISFSNLSSETTSQTTTILASTPGVKNL 107
 QY 61 OPTTVKNTNTTQOTOPSKPTTKORONKPPKPNNDHFEEVFNFPVCSICSNPTGMAIC 120
 DB 108 LPTTVKNTNTTQOTOPSKPTTKORONKPPKPNNDHFEEVFNFPVCSICSNPTGMAIC 167
 QY 121 KRIPNKKPGKKTTPKPKPTTKKDKLPQTTKPKKEVPTTKPEEPTINTTKNTTT 180
 DB 168 KRIPNKKPGKKTTPKPKPTTKKDKLPQTTKPKKEVPTTKPEEPTINTTKNTTT 227

Oy	181	LITNNTGANKLITSOMETFHSTSSNGCNLSPOSVFTSHPHQPSSPPTT	230
Dd	228	LITNNTGNPEHTSQETLHSTSSGPNPSQVYTTSETLSLPSSPNTT	277
 RESULT 7 O9YVB5 PRELIMINARY: PRT: 297 AA.			
ID	O9YVB5	AC	
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, Last sequence update)	
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)	
OC	Attachment glycoprotein G.		
OS	Human respiratory syncytial virus.		
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;		
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.		
OX	NCBI_TaxID=11250;		
RN	[1]	SEQUENCE FROM N.A.	
RC	STRAIN=MV2780;	MEDLINE=99022964; PubMed=9806017;	
RA	Sullender W.M., Mufson M.A., Prince G., Anderson L.J., Wertz G.W.;	"Antigenic and genetic diversity among the attachment proteins of group A respiratory syncytial viruses that have caused repeat infections in children."	
RL	J. Infect. Dis. 178:925-932(1998).		
DR	EMBL; AF065405; AAD02941.1; -		
DR	InterPro: IPR000925; Glycoprol.G.		
DR	InterPro: IPR003880; Pantane.attach.		
DR	InterPro: IPR002965; P.rich.extension.		
DR	Pfam: PF00802; Glycoprotein.G. 1.		
DR	PRINTS: PR01217; PRICHEXTENS.		
DR	PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.		
SQ	SEQUENCE 297 AA; 32572 MW; 5C10E719A05EF5C1 CRC64;		
 Query Match 86.8%; Score 1067; DB 12; Length 297; Best local Similarity 88.7%; Pred. No. 9, 4e-73; Matches 205; Conservative 7; Mismatches 19; Indels 0; Gaps 0.			
Oy	1	HKVTLTALIIDATSOIKNTPPYVLQDPLGLISFNSLETSQTTLILASTTPGVKSLN	60
Dd	67	HKVTSTTIITIQDATNOIKNTPPYLTQNOLGISPNSEITSLLITLDSTTPGVKSL	126
Oy	61	OPTVTCKNTTQOTQSPKRTKORONKPKNPNDFHFVFNFVPCSCSNPTCAIC	120
Dd	127	QSTTVGKNTTIQAOPSPTTKORONKPKSKRNDFHFVFNFVPCSCSNPTCAIC	186
Oy	121	KRIPIKKPGAKTKTKPKKEFKTKDKLKQGTKREVEFTTKTEBPTINTKTNIITT	180
Dd	187	KRIPIKKPGKRTTKTKPKRPTLKTTKKDPPKQTKTSKEVFTTKTEBPTINTKTNIITT	246
Oy	181	LITNNTGANKLITSOMETFHSTSSGCLSSOVSTSHESQSOSPNTTR	231
Dd	247	LITNNTIRNELITSOMETFHSTSSGPNPSQVYSITSEYSPSSPENTPR	297
 RESULT 8 O82066 PRELIMINARY: PRT: 297 AA.			
ID	O82066	AC	
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-MAR-2002	(TREMBLrel. 20, Last annotation update)	
DE	Glycoprotein.		
GN	G.		
OS	Human respiratory syncytial virus.		
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;		
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.		
OX	NCBI_TaxID=11250;		
RN	[1]	SEQUENCE FROM N.A.	
RP	STRAIN-SUBGROUP A;	MEDLINE=94335057; PubMed=8057427;	

[illegible]

Qy	61	QPTVTKRNTTITOTDQSPRTTKORÜKPNPKPNNDHEFEVFNFPQSIGSNBPTCAIC	120
Db	122	QSTVPRKNTMTTIOAPSCKPTTKORÜKPNPKPNNDHFEVFNFPQSIGSNBPTCAIC	181
Qy	121	KRIPNKKPGKTTTKPRPKPPFKTTKKDKLQÖTTKPREVPTRKEEPTINTTKNTTTT	180
Db	182	KRIPNKKPGKTTTKPRPKPPFKTTKKDKLPQÖTTKSEVPTRKEEPTINTTKNTTTT	241
Qy	181	LITNTNGNKLTQÖMETFHSTSEGNLSQÖVSTTSEHQSÖSSPNTTRQ	232
Db	242	LITNTNRNDELQÖMETFHSTSEGNLSQÖVSTTSEHQSÖSSPNTTRQ	293

RESULT 10	
Q9YVB4	
ID	Q9YVB4 PRELIMINARY; PRT; 298 AA.
AC	Q9YVB4;
DT	01-MAY-1999 (TREMBLrel. 10, Created)
DT	01-MAY-1999 (TREMBLrel. 10, last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE	Attachment glycoprotein G.
OS	Human respiratory syncytial virus.
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.
NC	NCBI_TaxID=11250;
LN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN-WV5222;
RX	MEDLINE=99022964; PubMed=9806017;
RA	Sullivan W.M., Wufson M.A., Prince G., Anderson L.J., Wertz G.W.,
RT	"Antigenic and genetic diversity among the attachment proteins of
RT	group A respiratory syncytial viruses that have caused repeat
RT	infections in children."
RL	J. Infect. Dis. 178:925-932(1998).
DR	EMBL; AF065406; AAD02942.1; -
DR	InterPro: IPR000925; Glycoprotei.G.
DR	InterPro: IPR003880; Ppatine.atlatch.
DR	Pfam: PF00802; Glycoprotein.G.1
DR	PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN_1.
Q0	SEQUENCE 298 AA: 32781 MW: A679E05547C5745 CRC64:

OC Virusess ssRNA negative-strand virusess; Mononegavirales;
OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
OX NCBI_TaxId=11250;
RN RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SUBGROUP A;
RX MEDLINE=94335057; PubMed=8057427;
RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabrisse S., Russi J.,
RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
RA Melero J.A.;
RT "Evolutionary pattern of human respiratory syncytial virus (subgroup
RT A): cocirculating lineages and correlation of genetic and antigenic
RT changes in the G glycoprotein.";
RL J. Virol. 68:5448-5459(1994).
DR EMBL: Z33432; CAA83875.1; ...
DR InterPro: IPR000925; Glycoprot-G.
DR InterPro: IPR003850; Pentate-attach.
DR Pfam: PF00802; Glycoprotein-G_1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
SO SEQUENCE 297 AA; 32551 MW; AD947B2318193306 CRC64;

RESULT	11
082071	
ID	Q82071
AC	Q82071.
DT	01-NOV-1996 (TREMBLrel_01, Created)
DT	01-NOV-1996 (TREMBLrel_01, Last sequence update)
DT	01-MAR-2002 (TREMBLrel_20, Last annotation update)
DE	Glycoprotein.
GN	G.
OS	Human respiratory syncytial virus.

Query Match	85.9%	Score 1056	DB 12	Length 297
Best Local Similarity	87.4%	Pred. No. 6,3e-72		
Matches 202	Conservative 9	Mismatches 20	Indels 0	Gaps 0
QY	1	HKVLTTLAIODATSOIKNTPTPTVLTODPOLGIFSNSLSEITSQTTLTILASTTPGVKSNL	60	
Db	67	HKVSTTTTIIQATNOIKNTPTPTVLTQNPOLGISNSLSEITSLLTILIDSTTPGVKSTL	126	
QY	61	OPTVTKNTNTTYYQVPSKPTTKQKQNNKPPNKNDFHEVEVNEVPCISCSNNPTCAIC	120	
Db	127	OSTVGINNTNTTQKQPKPTTKQKQNNKPPNKNDFHEVEVNEVPCISCSNNPTCAIC	186	
QY	121	KRIPNKKGGKKTPTTKPKPTTKKPKLPOPTTKPKPEVPTTKPPEEPTINTTKTNITTT	180	
Db	187	KRIPNKKGGKKTPTTKPKPTTKKPKLPOPTTKPKPEVPTTKPPEEPTINTTKTNITTT	246	
QY	181	LTNNNTGNPKLITSOMETFHSTSSGSLNSPQSVSTSHSPQSSPPMPTTR	231	
Db	247	PLTSTNTNPELTSOMETFHSTSSGSLNSPQSVSTSHSPQSSPPMPTTR	297	
RESULT 12				
Q906T3				
ID	Q906T3	PRELIMINARY	PRT	278 AA.
AC	Q906T3			
DT	01-MAY-2000 (TREMBlrel. 13, Created)			
DT	01-MAY-2000 (TREMBlrel. 13, Last sequence update)			
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)			
DE	Glycoprotein (Fragment).			
GN	G.			
OS	Human respiratory syncytial virus.			
OC	Viruses; ssRNA negative-strand viruses; Mononegavirales;			
OC	Paramyxoviridae; Pneumovirinae; Pneumovirus.			
OX	NCBI_TaxID=11250;			
RN	[1]			
RC	SEQUENCE FROM N.A.			
RC	STRAIN=93006;			
RX	MEDLINE=20283719; PubMed=10823752;			
RA	Choi E.H., Lee H.J.;			
RT	"Genetic diversity and molecular epidemiology of the G protein of			
RT	subgroups A and B of respiratory syncytial virus isolated over 9			
RT	consecutive epidemics in Korea"			
RL	J. Infect. Dis. 181:1547-1556(2000).			
DR	EMBL: AF19313; AAE2336.1; -			
DR	InterPro: IPR000925; Glycoprol.G.			
DR	InterPro: IPR003880; Pplantn_attnch.			
DR	InterPro: IPR002965; P_rich_extensn.			
DR	Pfam: PF00802; Glycoprotein.G; 1.			
DR	PRINTS: PR01217; PRICHEXTENS.			
DR	PROSITE: PS00012; PHOSPHOPANTHETHEINE; UNKNOWN.1.			
FT	NON_TER			
FT	1			
Q0	SEQUENCE 278 AA: 30443 MM: 8E2974F2E1AE15C9 CRC64:			

Query Match 85.8%; Score 1053; DB 12; Length 278;
 Best Local Similarity 87.4%; Pred. No. 7e-72;
 Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVLTALTIODATSOIKNTPTTYLTQDPLGISFNSLSEISQTTTILASTPGVKSNTL 60
 |||||
 DB 48 HKVLTALTIODATSOIKNTPTTYLTQDPLGISFNSLSEISQTTTILASTPGVKSNTL 107
 |||||
 QY 61 OPTTVKTKNTTTQTOPSKPTTKORONKPPNKNPNDFHEFEVNFVPCISCSNNPTCWAIC 120
 |||||
 DB 108 QSTTVKTKNTTTQTOPSKPTTKORONKPPNKNPNDFHEFEVNFVPCISCSNNPTCWAIC 167
 |||||
 QY 121 KRIPNKKPGKTTTKPTKPKPTTKKDKLPQTTKREVPPTKPEEPTINTKTNTT 180
 |||||
 DB 168 KRIPNKKPGKTTTKPTKPKPTTKKDKLPQTTKREVPPTKPEEPTINTKTNTT 227
 |||||
 QY 181 LITNNTGNPKLTSGOMETFHSSTSEGNLSPSOVSTSEHPSPSPNTT 230
 |||||
 DB 228 LITNNTGNPKLTSGOMETFHSSTSEGNLSPSOVSTSEHPSPSPNTT 277
 |||||

RESULT 13

Q9YVB2 PRELIMINARY; PRT; 292 AA.

AC Q9YVB2. (TREMBLREL. 10, Created)
 DT 01-MAY-1999 (TREMBLREL. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
 DE Attachment glycoprotein G (Fragment).
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN-MV1983;
 RA MEDLINE=99022964; PubMed=9806017;
 RT "Antigenic and genetic diversity among the attachment proteins of group A respiratory syncytial viruses that have caused repeat infections in children."
 RT J. Infect. Dis. 178:925-932(1998).
 RL EMBL: AF065408; AAD02944.1;
 DR InterPro: IPR000925; Glycoprol_G.
 DR InterPro: IPR003880; Panthe_attach.
 DR Pfam: PF00802; Glycoprotein_G; 1.
 DR PRINTS: PR01217; PRICHEXTENS.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 292 AA; 31964 MW; 8942A8DD0A402A4B CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 292;
 Best Local Similarity 87.9%; Pred. No. 1e-71;
 Matches 203; Conservative 8; Mismatches 20; Indels 0; Gaps 0;

QY 1 HKVLTALTIODATSOIKNTPTTYLTQDPLGISFNSLSEISQTTTILASTPGVKSNTL 60
 |||||
 DB 62 HKVLTSTTTIIONATNOIKNTPTTYLTQDPLGISFNSLSEISQTTTILASTPGVKSNTL 121
 |||||
 QY 61 OPTTVKTKNTTTQTOPSKPTTKORONKPPNKNPNDFHEFEVNFVPCISCSNNPTCWAIC 120
 |||||
 DB 122 QSTTVKTKNTTTQTOPSKPTTKORONKPPNKNPNDFHEFEVNFVPCISCSNNPTCWAIC 181
 |||||
 QY 121 KRIPNKKPGKTTTKPTKPKPTTKKDKLPQTTKREVPPTKPEEPTINTKTNTT 180
 |||||
 DB 182 KRIPNKKPGKTTTKPTKPKPTTKKDKLPQTTKREVPPTKPEEPTINTKTNTT 241
 |||||
 QY 181 LITNNTGNPKLTSGOMETFHSSTSEGNLSPSOVSTSEHPSPSPNTT 231
 |||||
 DB 242 LITNNTGNPKLTSGOMETFHSSTSEGNLSPSOVSTSEHPSPSPNTT 292
 |||||

RESULT 14

Q82058 PRELIMINARY; PRT; 297 AA.

ID Q82058
 AC Q82058.
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
 DE (MAD-1-89) subgroup A, G glycoprotein.
 GN G.
 OS Human respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=11250;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN-SUBGROUP A.
 RA MEDLINE=94335057; PubMed=8057427;
 RA Garcia O., Martin M., Dopazo J., Arbizu J., Fabris S., Russi J.,
 RA Hortal M., Perez-Brena P., Martinez I., Garcia-Barreno B.,
 RA Melero J.A.;
 RT "Evolutionary pattern of human respiratory syncytial virus (subgroup A): cocirculating lineages and correlation of genetic and antigenic changes in the G glycoprotein."
 RT J. Virol. 68:5448-5459(1994).
 RL EMBL: Z33456; CAA83879.1;
 DR InterPro: IPR000925; Glycoprol_G.
 DR InterPro: IPR003880; Panthe_attach.
 DR Pfam: PF00802; Glycoprotein_G; 1.
 DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
 SQ SEQUENCE 297 AA; 32741 MW; 0E567A174F64964 CRC64;

Query Match 85.7%; Score 1053; DB 12; Length 297;
 Best Local Similarity 87.4%; Pred. No. 1.1e-71;
 Matches 201; Conservative 8; Mismatches 21; Indels 0; Gaps 0;

QY 1 HKVLTALTIODATSOIKNTPTTYLTQDPLGISFNSLSEISQTTTILASTPGVKSNTL 60
 |||||
 DB 67 HKVLTALTIODATSOIKNTPTTYLTQDPLGISFNSLSEISQTTTILASTPGVKSNTL 126
 |||||
 QY 61 OPTTVKTKNTTTQTOPSKPTTKORONKPPNKNPNDFHEFEVNFVPCISCSNNPTCWAIC 120
 |||||
 DB 127 QSTTVKTKNTTTQTOPSKPTTKORONKPPNKNPNDFHEFEVNFVPCISCSNNPTCWAIC 186
 |||||
 QY 121 KRIPNKKPGKTTTKPTKPKPTTKKDKLPQTTKREVPPTKPEEPTINTKTNTT 180
 |||||
 DB 187 KRIPNKKPGKTTTKPTKPKPTTKKDKLPQTTKREVPPTKPEEPTINTKTNTT 246
 |||||
 QY 181 LITNNTGNPKLTSGOMETFHSSTSEGNLSPSOVSTSEHPSPSPNTT 230
 |||||
 DB 247 LITNNTGNPKLTSGOMETFHSSTSEGNLSPSOVSTSEHPSPSPNTT 296
 |||||

RESULT 15

Q86359 PRELIMINARY; PRT; 295 AA.

ID Q86359
 AC Q86359.
 DT 01-NOV-1996 (TREMBLREL. 01, Created)
 DT 01-NOV-1996 (TREMBLREL. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLREL. 20, Last annotation update)
 DE G protein (Fragment).
 OS respiratory syncytial virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Pneumovirinae; Pneumovirus.
 OX NCBI_TaxID=12814;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC STRAIN-RSB89-6256;
 RA MEDLINE=91374005; PubMed=1895054;
 RA Cane P.A., Matthews D.A., Pringle C.R.;
 RT "Identification of variable domains of the attachment (G) protein of subgroup A respiratory syncytial viruses."
 RT J. Gen. Virol. 72:2091-2096(1991).

```
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-RSB89-6256;
RA Cane P.A.;
RL Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR EMBL: X73353; CA51764.1; -.
DR InterPro: IPR000925; Glycoprot.G.
DR InterPro: IPR003880; Pnanine_atlch.
DR Pfam: PF00802; Glycoprotein_G; 1.
DR PROSITE: PS00012; PHOSPHOPANTHEINE; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 32490 MW; 48D835F670FF8006 CRC64;

Query Match      85.6%; Score 1052; DB 12; Length 295;
Best Local Similarity 87.4%; Pred. No. 1.3e-71;
Matches 201; Conservative 6; Mismatches 23; Indels 0; Gaps 0;

QY 1 HKVTLTTAIIODATSOIKNTTPYLTODPOLGISFSNLSSEITISQTTTIIASTTPGVKSNL 60
   |||||
Db 65 HKVTLTTAIIODATSOIKNTTPYLTIONPOLGISFSNLSSETTSQPTTPAPTTPSAESTP 124
QY 61 OPTTVTKNTTTTQTOPSKPTTKORONKPPNKPNDHFVFNVPFCISNNPTCWAIC 120
   |||||
Db 125 QSTTVKTKNTTTTQIOPSKPTTKORONKPPNKPNDHFVFNVPFCISNNPTCWAIC 184
QY 121 KRIPNKKPGKKTTPPKPKPTKTKKDLKPOTTKPEVPTKPTKEPTINTKTNITPT 180
   |||||
Db 185 KRIPNKKPGKKTTPPKPKPTKTKKDLKPOTTKPEVPTKPTKEPTINTKTNIRTT 244
QY 181 LLTNNTGNPKLTSSOMETFSSTSEGNLSPOYSTSEHSPQSSPPTT 230
   |||||
Db 245 LTTNTTGNPEYTSQKEITLHSTSPGNSPSPQYTTSEYSPQSPSPSNT 294
```

Search completed: May 1, 2003, 13:46:24
Job time : 29.0151 secs

